Building Resilient Enterprise Campus Networks

BRKCRS-3032

Presenter

- Bryan Osoro – CCIE #8548 (R&S/Sec/SP/Voice)
  - Technical Solutions Architect
  - Worldwide Data Center & Virtualization

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Confidential

What Are Your Uptime Requirements?

- Campus network design is evolving in response to multiple drivers
  - User Expectations: Always ON Access to communications
  - Vertical Requirements: Financial, Healthcare, 7x24x365 Global access
  - Technology Requirements: Unified Communications
  - Unexpected Requirements: Worms, Viruses, …

Requires a Structured 'and' Resilient Design

Global Enterprise Availability

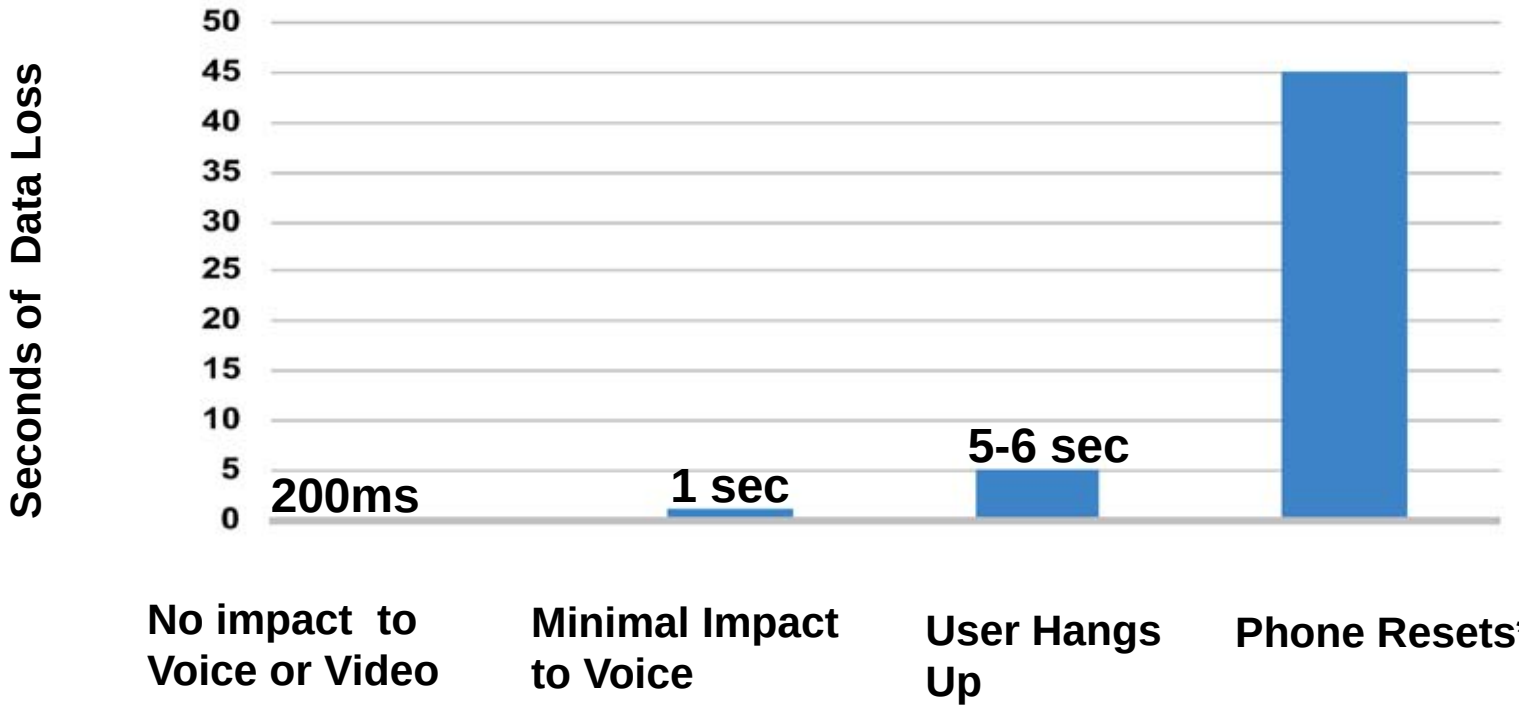
Collaboration and Real-Time Communication

Security

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

How Does Downtime Affect Voice or Video?

- Availability Requirements for UC are more than just five 9's
- Also need to consider the subjective impact to real time communications



*** The time for a phone to reset is variable and depends on the signaling protocol (SCCP or SIP) and the state of the call (active, ringing, …)**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

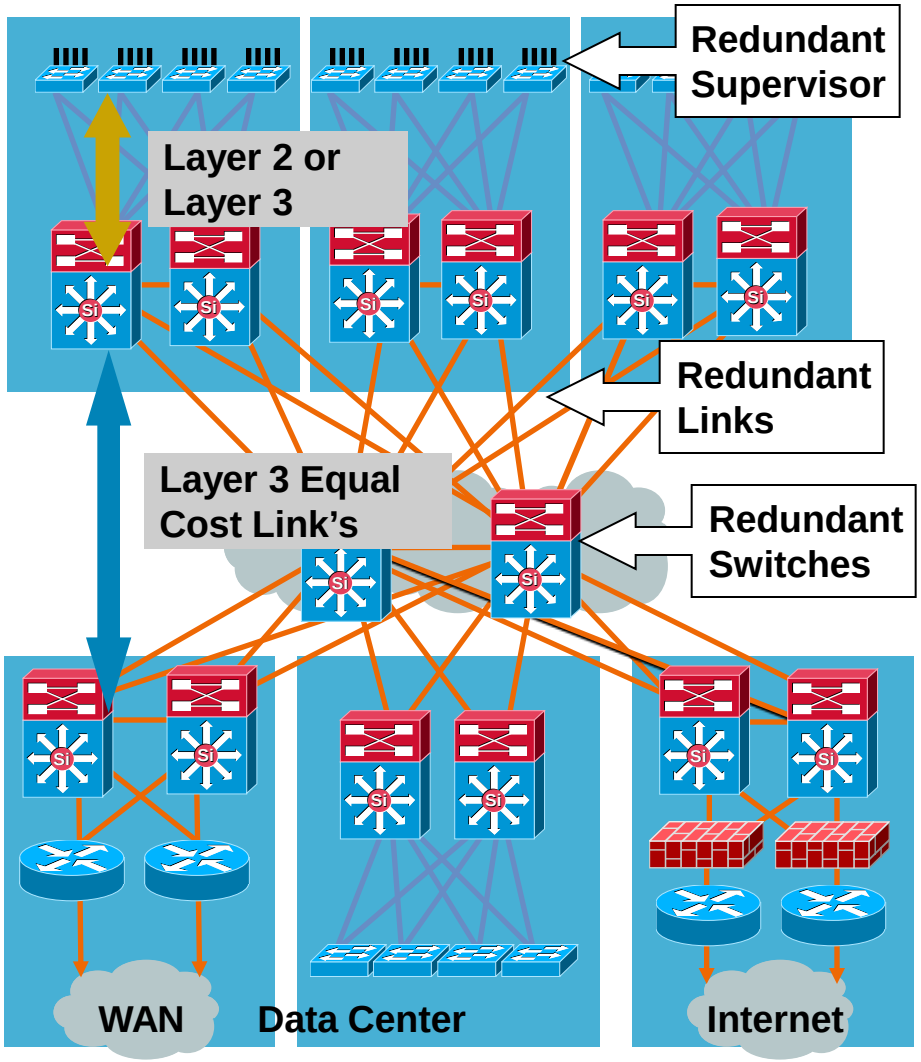
High Availability Campus Design
Agenda

- **Network Level Resiliency**
  - High Availability Design Principles
  - Redundancy in the Distribution Block
  - Campus Routing Best Practices
- **System Level Resiliency**
  - Integrated Hardware and Software Resiliency
    - **NSF/SSO**
    - **ISSU**
  - System Management Resiliency
    - **GOLD & EEM**
- **Hardening the Campus Network Design**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Structured and Modular Designs Work Best

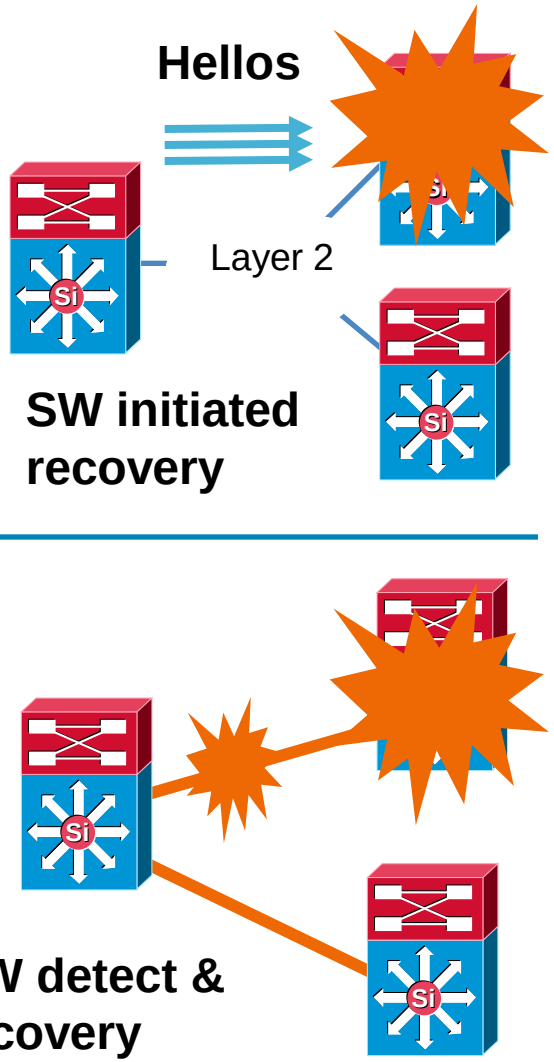
- Optimize the interaction of the physical redundancy with the network protocols
  - Provide the necessary amount of redundancy
  - Pick the right protocol for the requirement
  - Optimize the tuning of the protocol
- The network looks like this so that we can map the protocols onto the physical topology
- We want to build networks that look like this



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Be Aware of Extending Direct Links with a Passive Switch

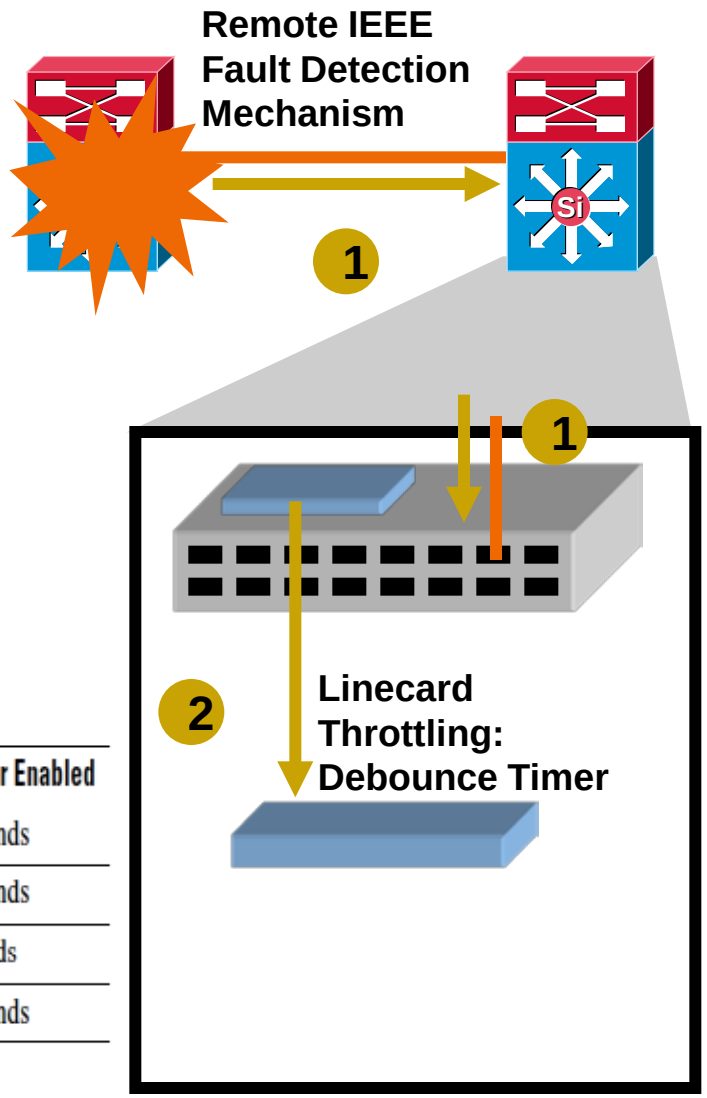
- Indirect link failures are harder to detect
- Caused by switches not participating in the recovery protocol such as hubs or dumb switches
- Bi-directional Forwarding Detection (BFD) provides a low-overhead method of detecting failures between two routing peers
- With no direct HW notification of link loss or topology change convergence times are dependent on SW notification via Spanning Tree BPDUs or Routing Protocol Hellos



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Did You Know that Copper Takes Longer than Fiber for Failure Detection?

- Direct point to point fiber provides for fast failure detection
- IEEE 802.3z and 802.3ae link negotiation define the use of **Remote Fault Indicator & Link Fault Signaling mechanisms**
- Debounce (Catalyst default to disabled)
- Carrier-delay = 0



Port Type	Debounce Timer Disabled	Debounce Timer Enabled
Ports operating at 10 Mpbs or 100 Mbps	300 milliseconds	3100 milliseconds
Ports operating at 1000 Mpbs or 10 Gbps over copper media	300 milliseconds	3100 milliseconds
Ports operating at 1000 Mpbs or 10 Gbps over fiber media	10 milliseconds[1]	100 milliseconds
WS-X6502-10GE 10-Gigabit ports	1000 milliseconds	3100 milliseconds

1. 10 milliseconds with Release 12.2(18)SXF13 and later releases.

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Routed Interfaces Offer Better Convergence Properties than SVIs

- Configuring L3 routed interfaces provides for faster convergence than a L2 switchport with an associated L3 SVI

~ 8 msec loss

1. Link Down
2. Interface Down
3. Routing Update

```
21:38:37.042 UTC: %LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet3/1, changed state to down
21:38:37.050 UTC: %LINK-3-UPDOWN: Interface GigabitEthernet3/1, changed state to down
21:38:37.050 UTC: IP-EIGRP(Default-IP-Routing-Table:100): Callback: route_adjust GigabitEthernet3/1
```

~ 150-200 msec loss

1. Link Down
2. Interface Down
3. Autostate
4. SVI Down
5. Routing Update

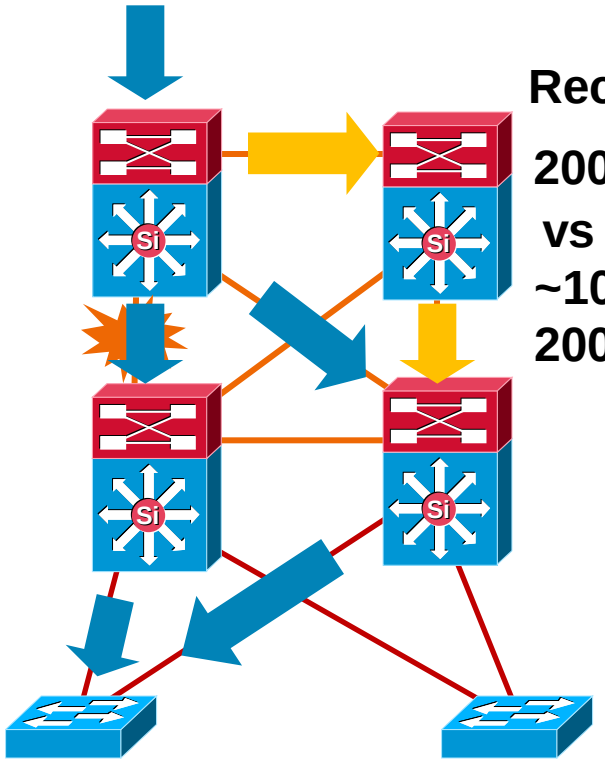
```
21:32:47.813 UTC: %LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet2/1, changed state to down
21:32:47.821 UTC: %LINK-3-UPDOWN: Interface GigabitEthernet2/1, changed state to down
21:32:48.069 UTC: %LINK-3-UPDOWN: Interface Vlan301, changed state to down
21:32:48.069 UTC: IP-EIGRP(Default-IP-Routing-Table:100): Callback: route, adjust Vlan301
```


© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Build Triangles Instead of Squares

- This design enables effective use of Equal Cost Multi-Path or Multi-chassis Etherchannel
- Time to restore traffic flows is based on
  - Time to detect link failure
  - Update the HW forwarding
- *No dependence on external events* (no routing protocol or spanning tree convergence required)
- Increased delay could be added if an alternate router has to take over in a layer 2 to the access design

Equal Cost Links: Link/Box Failure Does Not Require Multi-Box Interaction



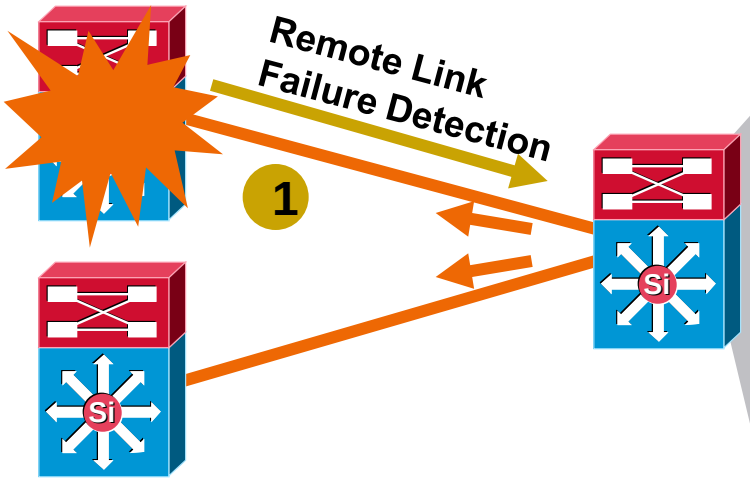
Recovery

200ms vs ~1000 to 2000ms

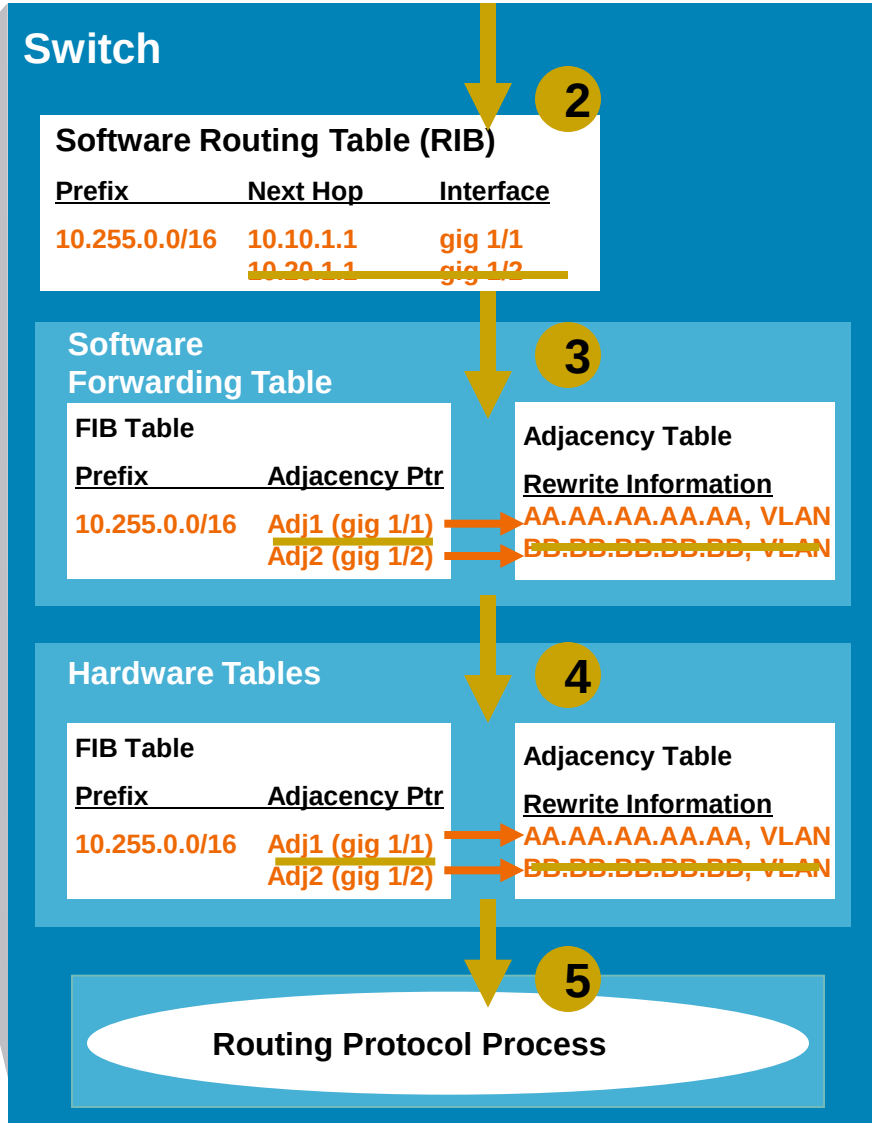
All Links Forwarding: In an Environment with All Links Active Traffic Is Restored Based on HW Recovery

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Routing Failover Takes Only 5 Steps



1. Link failure detection
2. Removal of the entries in the routing table
3. Update of the software CEF table to reflect to loss of the next hop adjacencies
4. Update of the hardware tables
5. Routing protocol notification and reconvergence



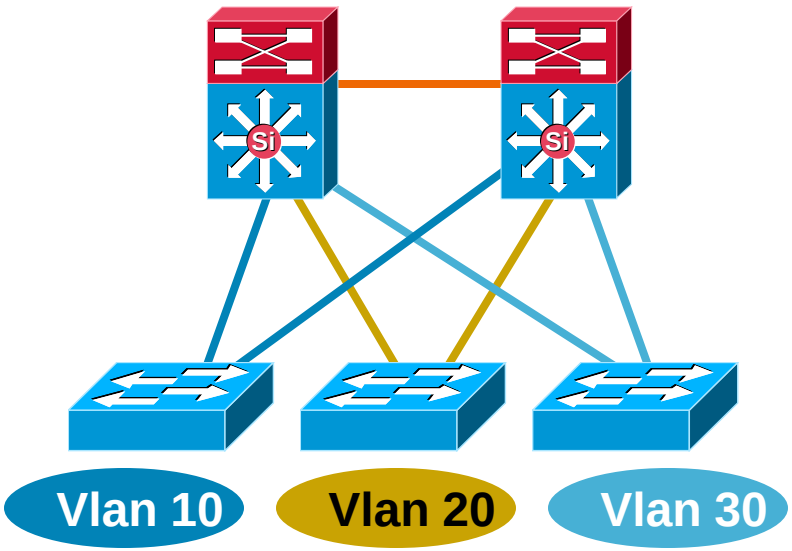
© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Confidential

High Availability Campus Design
Agenda

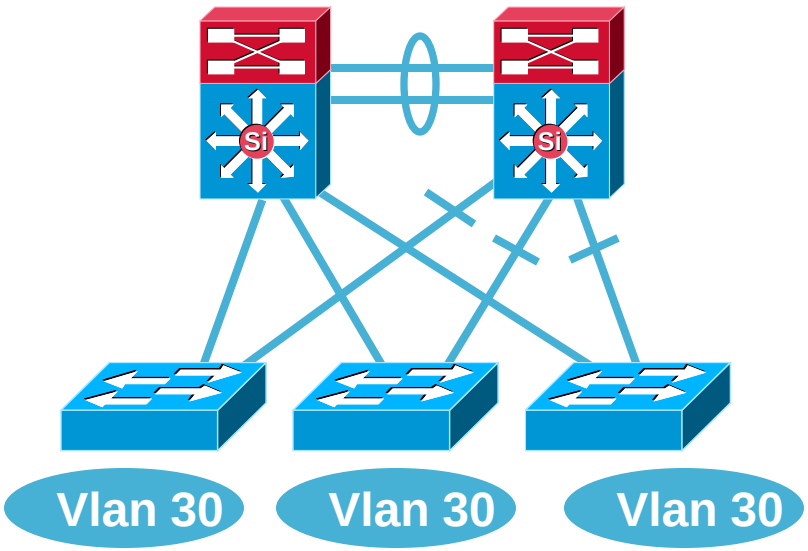
- **Network Level Resiliency**
  - High Availability Design Principles
  - Redundancy in the Distribution Block
  - Campus Routing Best Practices
- **System Level Resiliency**
  - Integrated Hardware and Software Resiliency
    - **NSF/SSO**
    - **ISSU**
  - System Management Resiliency
    - **GOLD & EEM**
- **Hardening the Campus Network Design**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

We Will Be Talking About Solutions for Two Distribution Block Models



- Each access switch has unique VLAN's
- No layer 2 loops
- Layer 3 link between distribution
- No blocked links

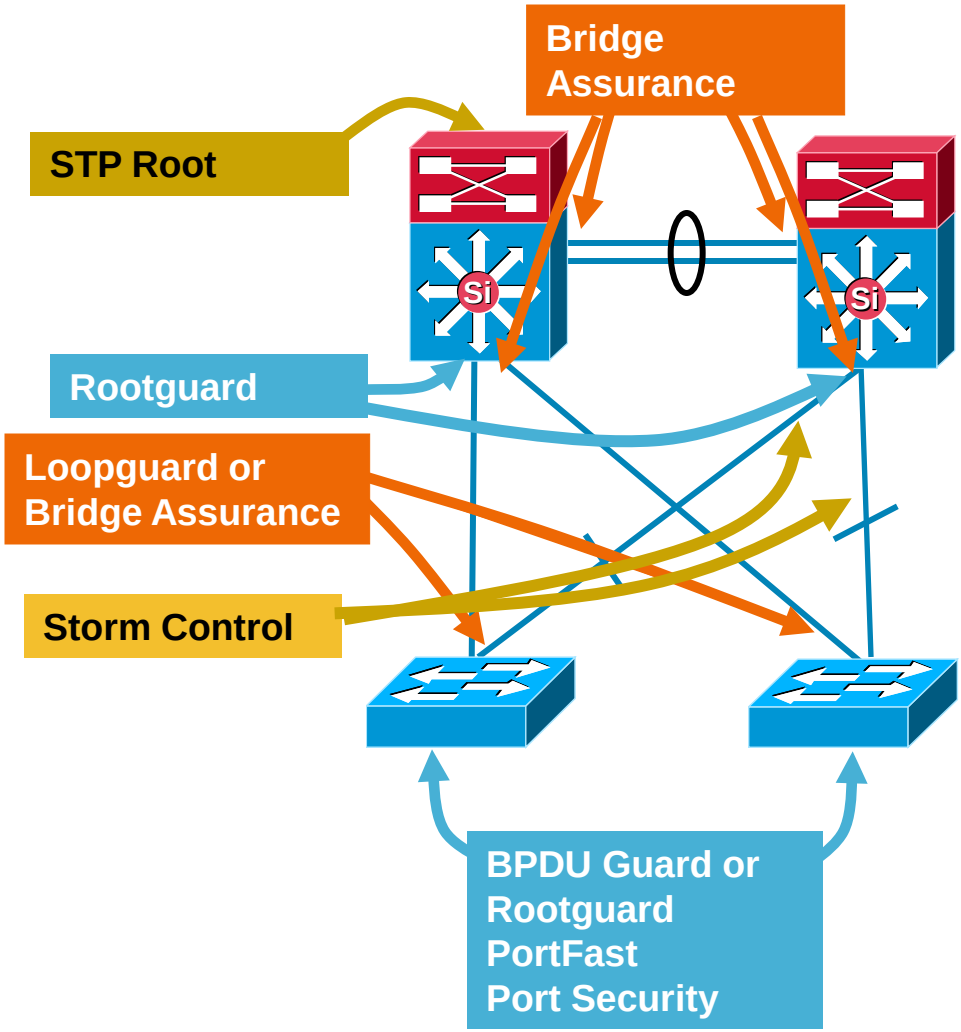


- At least some VLAN's span multiple access switches
- Layer 2 loops
- Layer 2 and 3 running over link between distribution
- Blocked links

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

If Spanning Tree Is Used Multiple Features Are Needed to Protect From Failure

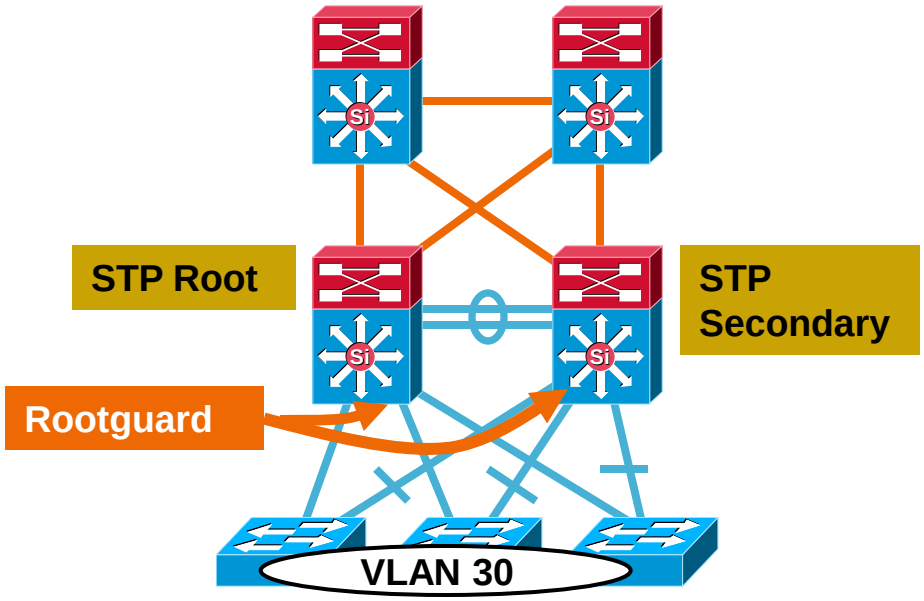
- Place the root bridge in the distribution layer for optimal traffic flow
- Some hardware issues can cause spanning tree to fail
- Understand and optimize the L2 element of your design
- L2 has no native mechanism to dampen a problem
- Storm control can be utilized on backup links to limit Broadcast traffic volumes during a STP Loop
  - Utilize Sup720 rate limiters or SupIV/V/6E with HW queuing structure



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Root Guard Prevents Root from Unexpectedly Moving

- Enable Root Guard on links connecting to access layer to protect from edge switches becoming root and causing sub-optimal traffic flow
- Forces Layer 2 LAN interface to be a designated port. If port receives a superior BPDU, Root Guard puts the interface into the root-inconsistent (blocked) state
- Channel the trunk between Distribution Switches so failure doesn't break topology

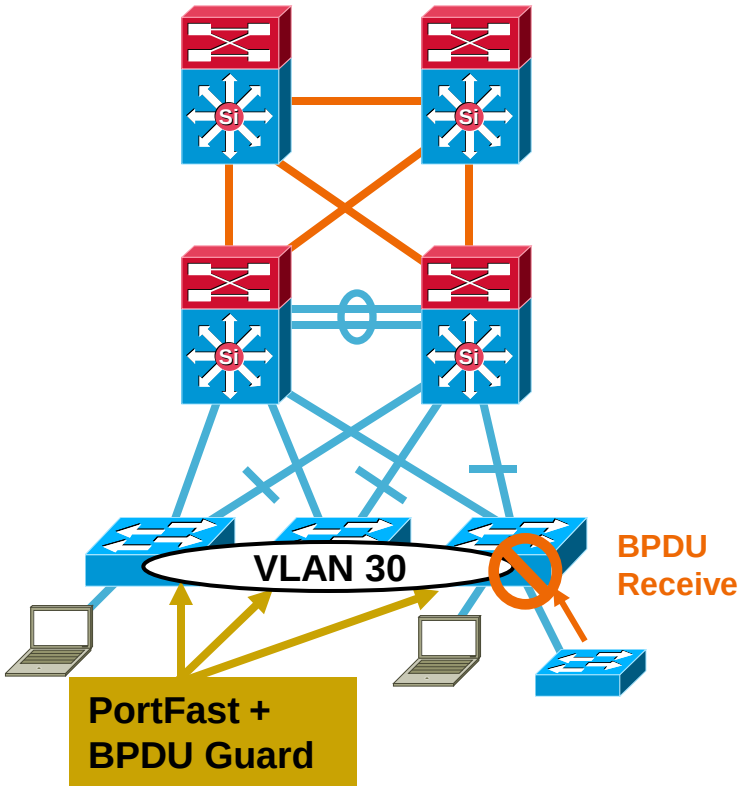



```
Router(config-if)# switchport
Router(config-if)# spanning-tree guard root

%SPANTREE-2-ROOTGUARDBLOCK: Port 3/3 tried to become non-designated in VLAN 800.
Moved to root-inconsistent state
```


© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

PortFast and BPDU Guard Protects Access Switch Ports

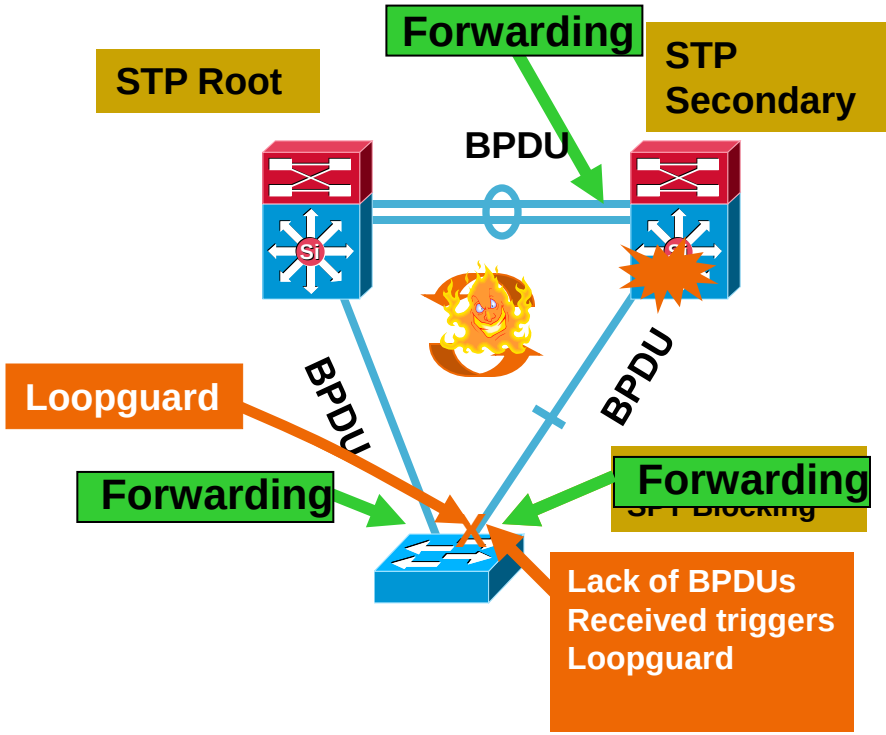
- Enable PortFast on edge ports to allow them to quickly move to forwarding bypassing listening and learning and avoid TCN (RPVST Topology Change Notification) messages
- BPDU Guard prevents ports configured with PortFast from being incorrectly connected to another switch
- Enable BPDU Guard to prevent loops by moving PortFast-configured interfaces that receive BPDUs to errdisable state
- When enabled globally, BPDU Guard applies to all interfaces that are in an operational PortFast state




```
Router(config-if)#spanning-tree portfast
Router(config)#spanning-tree portfast edge default
Router(config-if)#spanning-tree bpduguard enable
Router(config)#spanning-tree portfast edge bpduguard default

1w2d: %SPANTREE-2-BLOCK_BPDUGUARD: Received BPDU on port FastEthernet3/1 with
BPDU Guard enabled. Disabling port.
1w2d: %PM-4-ERR_DISABLE: bpduguard error detected on Fa3/1, putting Fa3/1 in
err-disable state
```

Loop Guard Protects Network from Unidirectional Links and Software Errors

- BPDUs flow between switches providing a means to detect a loop and block the higher cost path to the root
- If BPDUs stop flowing due to soft errors, max-age timer expires, port transitions from blocked to listening then forwarding state creating a loop
- When Loop Guard is enabled and BPDUs are no longer received on a non-Designated port the port is moved to a STP loop-inconsistent blocking state preventing a loop
- Works in PVST, RPVST+, and MST STP Domains



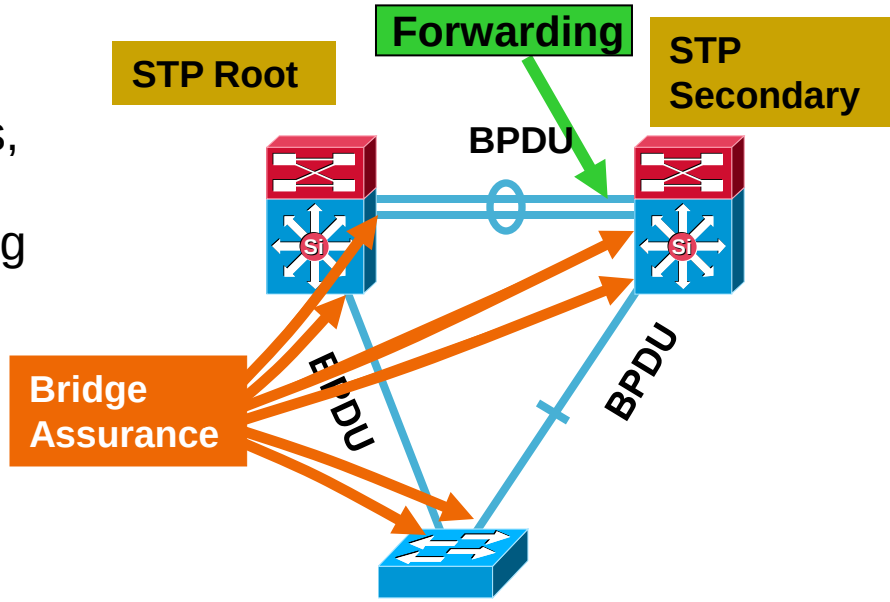

```
Router(config)# interface gigabitEthernet 2/1
Router(config-if)# spanning-tree guard loop
Router(config)# spanning-tree loopguard default

%SPANTREE-2-LOOPGUARD_BLOCK: Loop guard blocking port gigabitEthernet 2/1 on VLAN0050
```

The port is moved to a STP loop-inconsistent blocking state preventing a loop
Automatic recovery upon receiving BPDUs on port

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Bridge Assurance Also Protects Net from Unidirectional Links and Software Errors

- BPDUs flow between switches providing a means to detect a loop and block the higher cost path to the root
- If BPDUs stop flowing due to soft errors, max-age timer expires, port transitions from blocked to listening then forwarding state creating a loop
- Bridge Assurance is recommended network wide with latest software
- Point-to-point links only
- Only works in RPVST+, and MST STP Domains

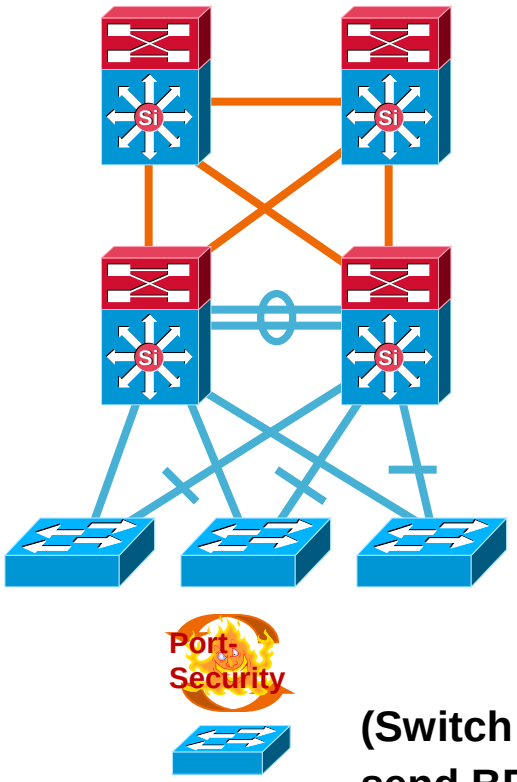



```
%STP-2-BRIDGE_ASSURANCE_BLOCK: Bridge Assurance blocking port Ethernet 2/48 VLAN0700
Switch# sh spanning vlan 700 | in –i bkn
Eth2/48             Altn BKN*4             l28.304    Network P2p *BA_Inc
Switch#
```

The port is moved to a STP loop-inconsistent blocking state preventing a loop
Automatic recovery upon receiving BPDUs on port

Dynamic Port Security Mac Move Violation Watches for Access Layer Loops

- Some consumer grade ‘personal’ switches don’t send BPDU’s
- If there are no BPDU’s BPDU Guard can’t detect and ‘stop’ them
- If no BPDU’s STP can’t prevent/stop a loop at the edge
- Port security detects MAC learned address on 2 ports and error disables or ignores traffic on second interface breaking the loop



(Switch that doesn’t send BPDU’s)

```
switchport port-security
switchport port-security maximum 10
switchport port-security violation shutdown
switchport port-security aging time 2
switchport port-security aging type inactivity
```

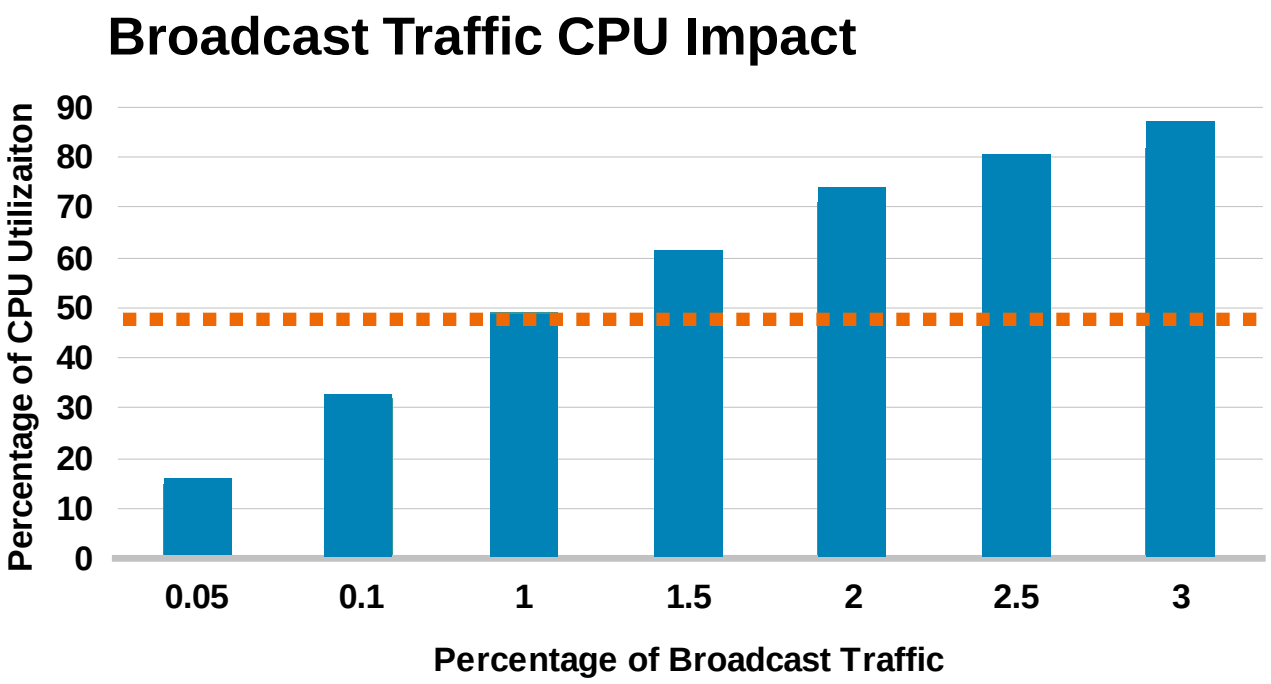
If Violation is set for Error-Disable, the Following Log Message Will Be Produced:

4w6d: %PM-4-ERR_DISABLE: Psecure-Violation Error Detected on Gi3/2, Putting Gi3/2 in Err-Disable State

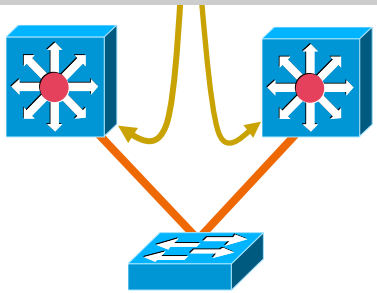
Harden the Network Links

Storm Control

- **Protect the network from intentional and unintentional flood attacks e.g. STP loop**
- **Limit the combined rate of broadcast and multicast traffic to normal peak loads**
- **Limit broadcast and when possible multicast to 1.0% of a GigE link to ensure distribution CPU stays in safe zone**



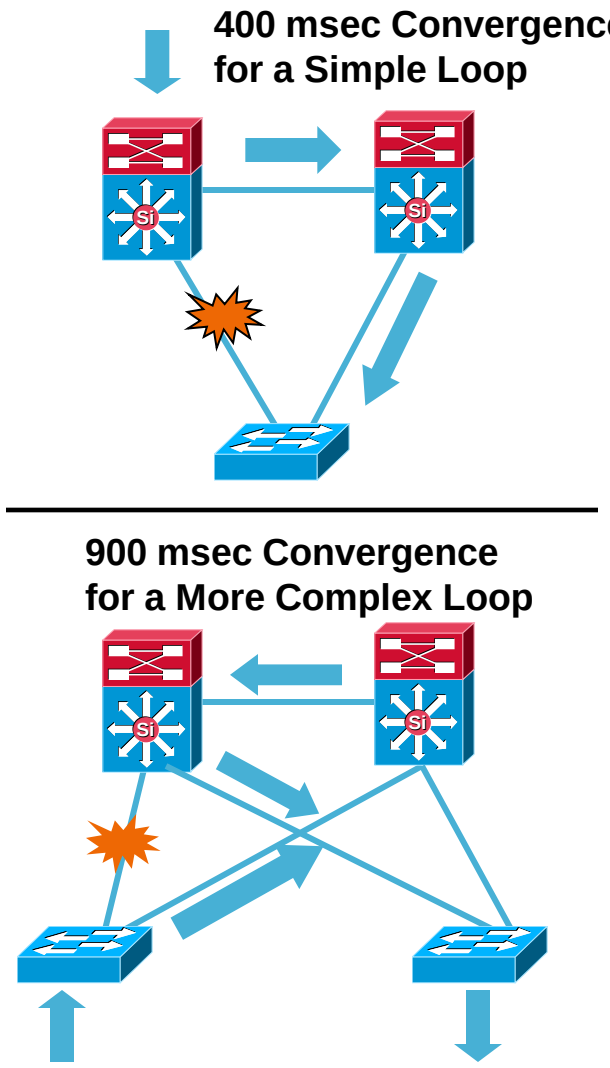

```
! Enable storm control

storm-control broadcast
level 1.0
storm-control multicast
level 1.0
```



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Convergence Is Longer with Complex Spanning Tree Loops and Older Protocols

- Time to converge is dependent on the protocol implemented 802.1d, 802.1s or 802.1w (all now a part of IEEE 802.1d 2004 spec)
- It is also dependent on:
  - Size and shape of the L2 topology (how deep is the tree)
  - Number of VLAN's being trunked across each link
  - Number of ports in the VLAN on each switch
- Complex Topologies Take Longer to Converge
- Multi-chassis Etherchannel is a good option to avoid this increase in convergence



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Even with Faster Convergence from RPVST+ We Still Have to Wait on FHRP Convergence

VRRP Config

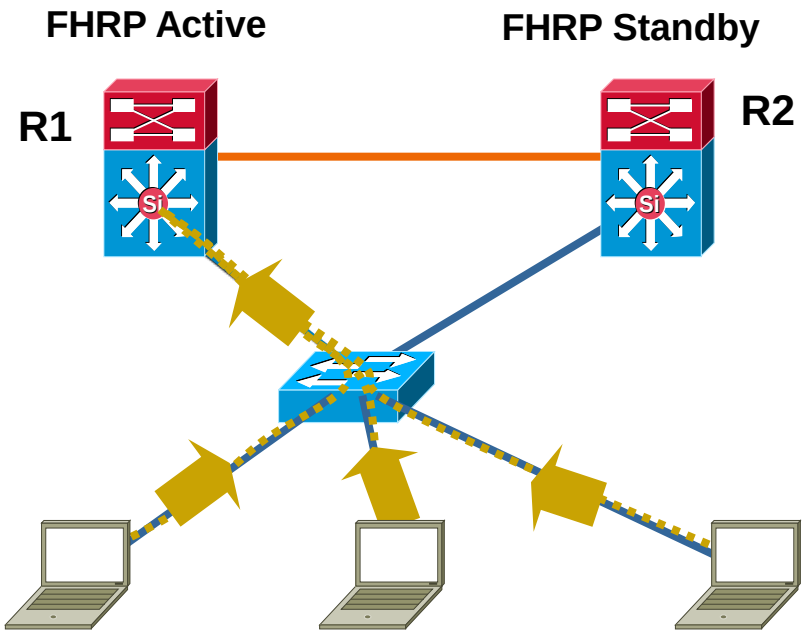
```
interface Vlan4
 ip address 10.120.4.1 255.255.255.0
 ip helper-address 10.121.0.5
 no ip redirects
 vrrp 1 description Master VRRP
 vrrp 1 ip 10.120.4.1
 vrrp 1 timers advertise msec 250
 vrrp 1 preempt delay minimum 180
```

HSRP Config

```
interface Vlan4
 ip address 10.120.4.2 255.255.255.0
 standby 1 ip 10.120.4.1
 standby 1 timers msec 250 msec 750
 standby 1 priority 150
 standby 1 preempt
 standby 1 preempt delay minimum 180
```

GLBP Config

```
interface Vlan4
 ip address 10.120.4.2 255.255.255.0
 glbp 1 ip 10.120.4.1
 glbp 1 timers msec 250 msec 750
 glbp 1 priority 150
 glbp 1 preempt
 glbp 1 preempt delay minimum 180
```

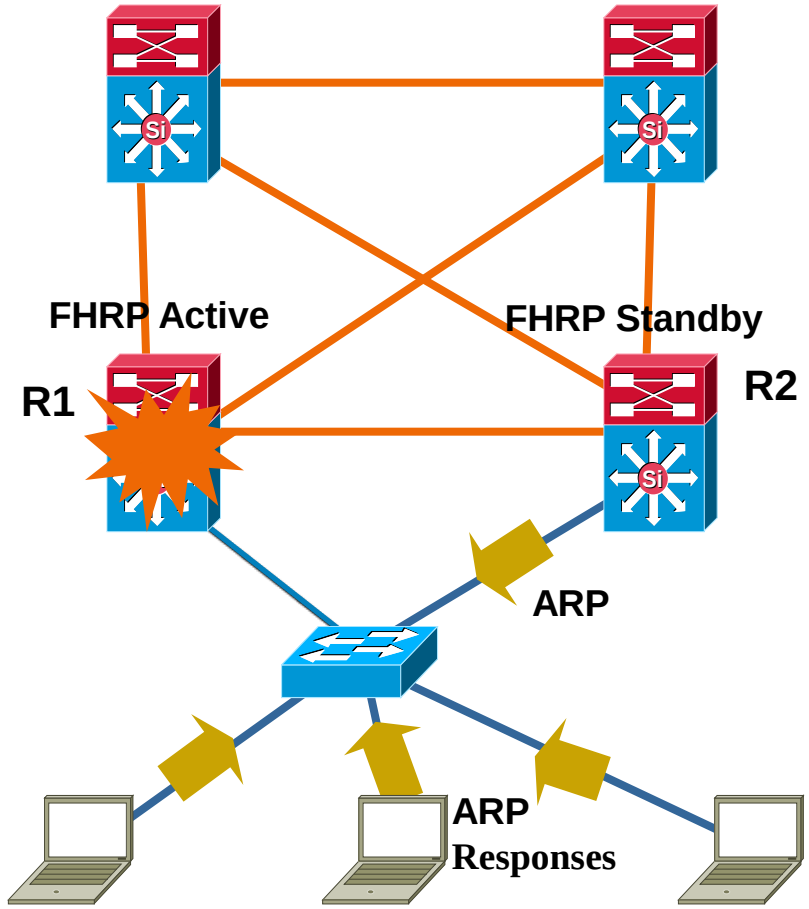


- GLBP offers load balancing within a VLAN
- For Voice, sub-second Hello timer enables < 1 Sec traffic recovery upstream
- Not necessary for Multi-chassis Etherchannel designs

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

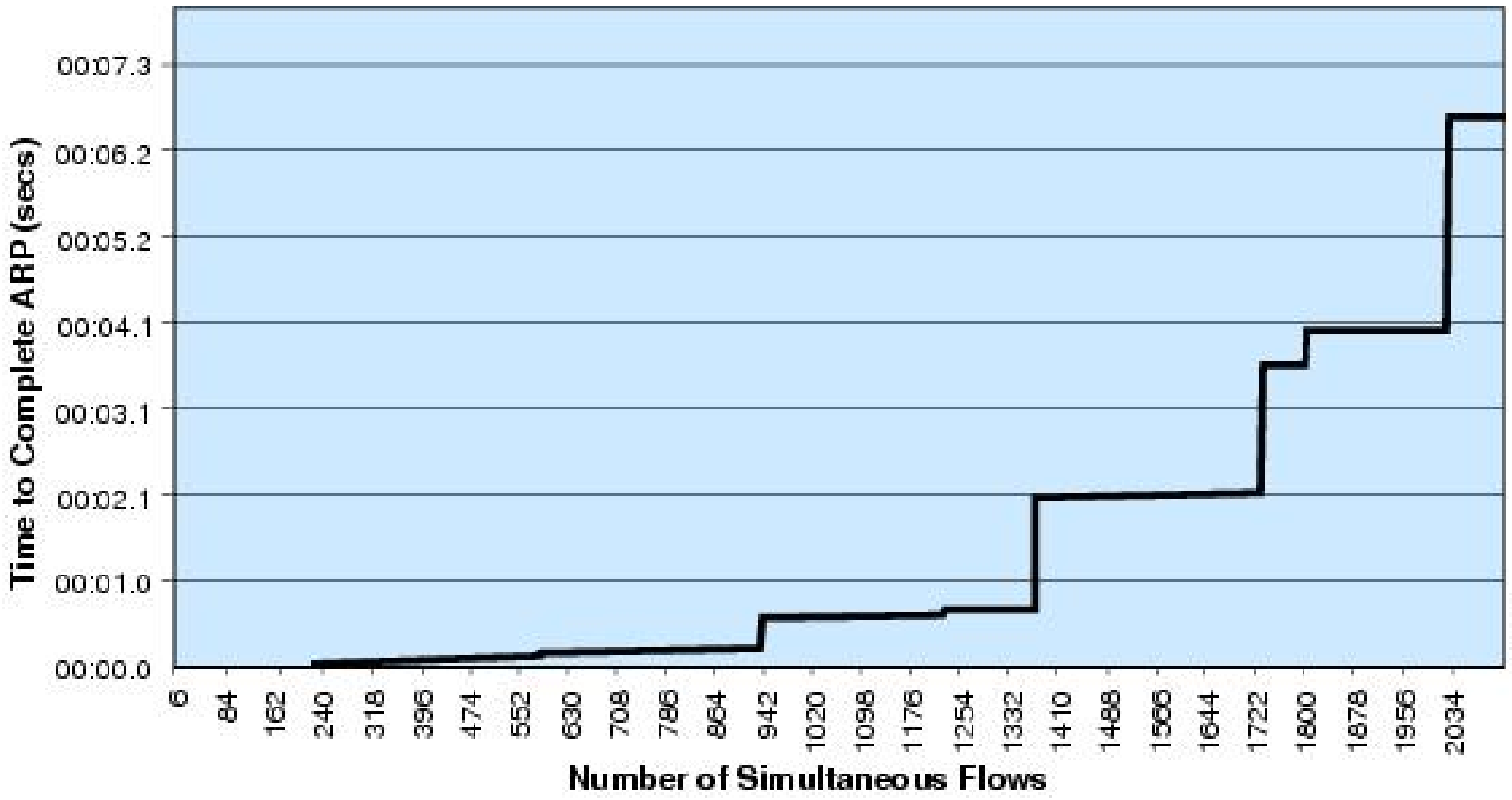
Inbound Traffic Has to Wait on ARP Resolution

- The ARP process is a CPU bound process which is rate limited to prevent Denial of service Attacks.
- Does not apply to Multi-Chassis Etherchannel designs



© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Confidential

Layer 3 Boundary Should Be as Close to Edge as Possible to Minimize ARP Processing Delay

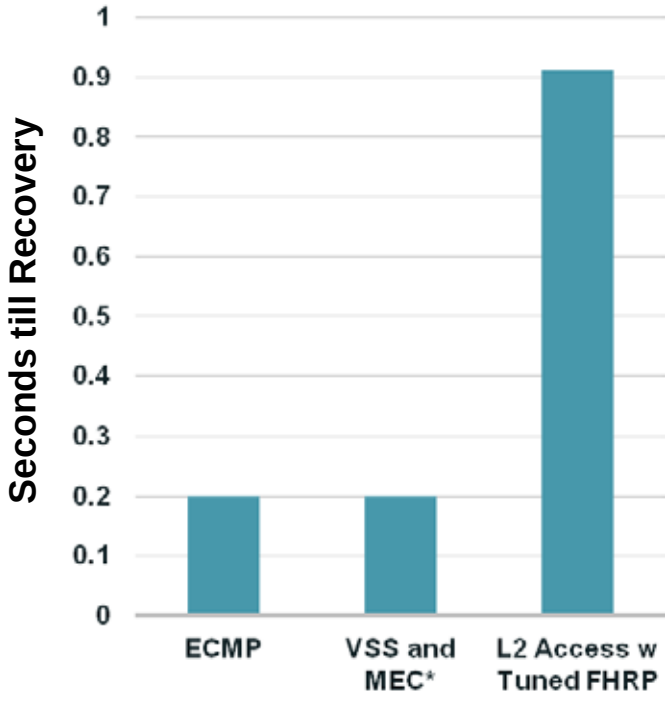


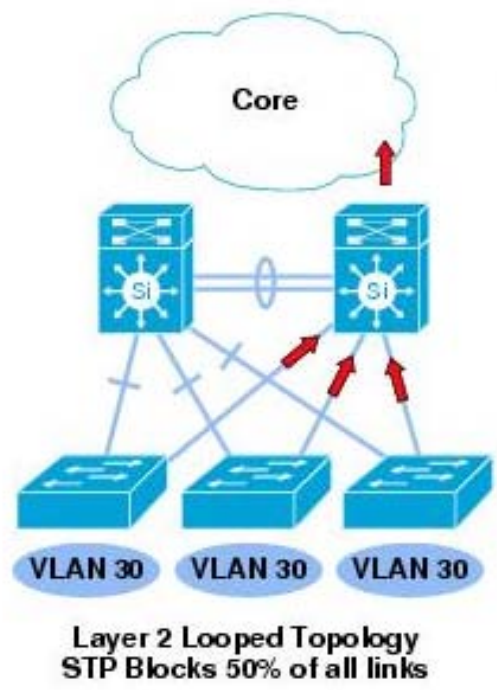
When an alternate router has to take over for inbound traffic, there is a delay that is based on the number of active flows.

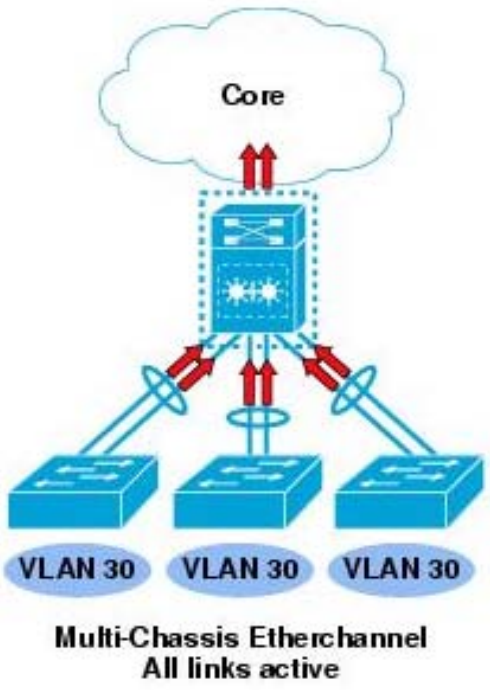
© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Multi-Chassis EtherChannel Solutions as an Alternative to STP and FHRP

- Spanning tree loops are replaced with ether-channel connections
- FHRP nodes are virtualized and ARP cache is replicated
- All links are fully utilized based on Ether-channel load balancing
- Similar benefits to L3 Equal Cost Multi Path





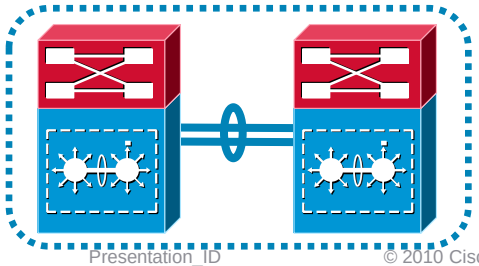


All rights reserved. Cisco Confidential

Three Options for Multi-Chassis EtherChannel Designs to Remove Spanning Tree

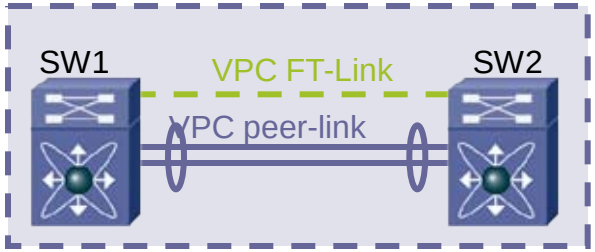
Virtual Switching System

- Single control plane
- Single management plane
- Single supervisor per chassis
- Automatic port config sync (single control plane)
- Single L3 domain (single SVI) **no need for FHRP**



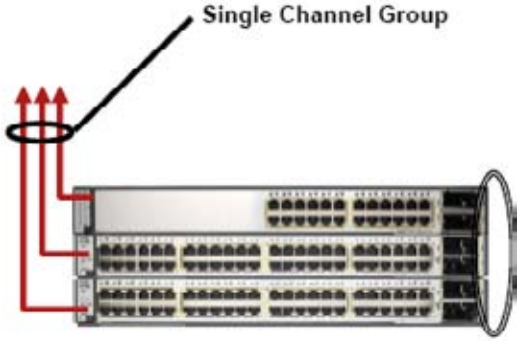
Virtual Port Channel

- Separate control plane
- Separate management plane with VPC state synchronization (CFS)
- Redundant supervisors per chassis with hitless SSO
- Manual port sync config (DataCenterNetworkMgr)
- Local SVI **HSRP/PIM forwarding enhancements to act as active-active pair**



Stackwise+

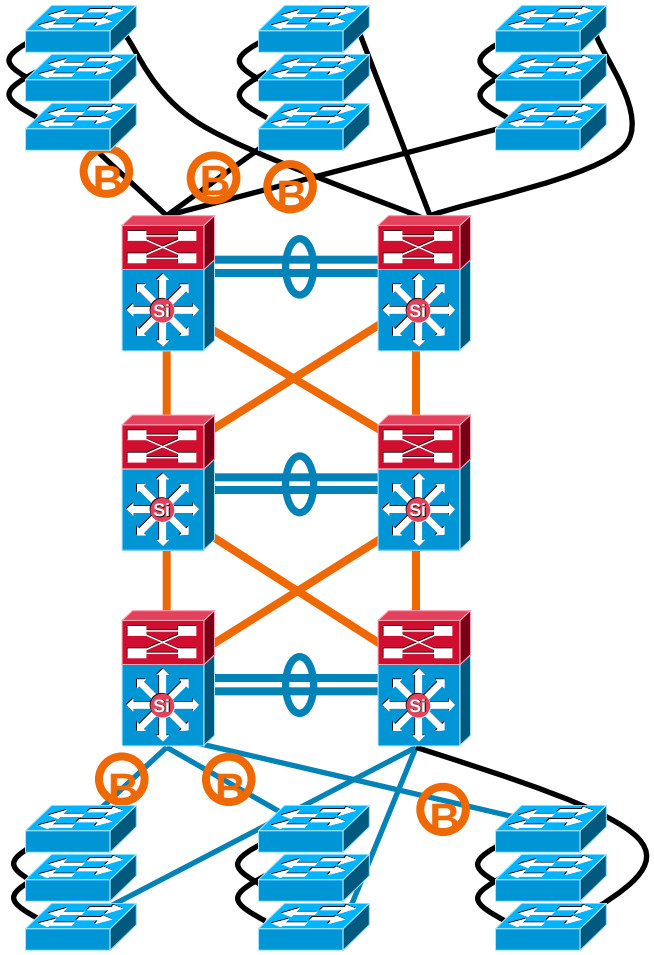
- Single control plane controlled by Master Switch
- Master switch controls etherchannel
- Redundant master switches per stack
- Automatic port config sync (single control plane)
- Stack appears as a single router, **no need for FHRP**



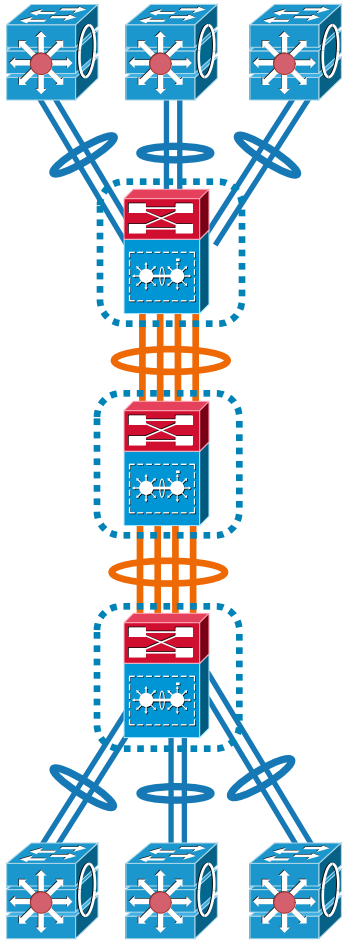
© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

VSS Enabled Campus Design

End-to-End VSS Design Option



STP-Based
Redundant Topology

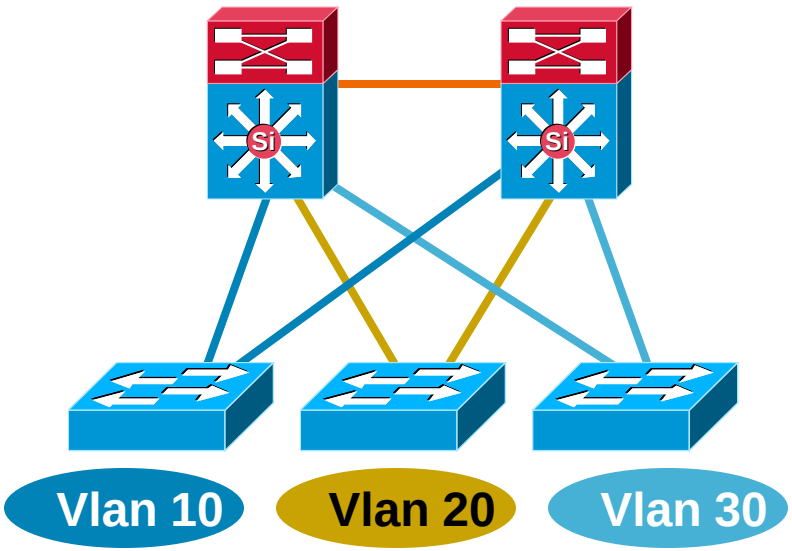


Fully Redundant
Virtual Switch Topology

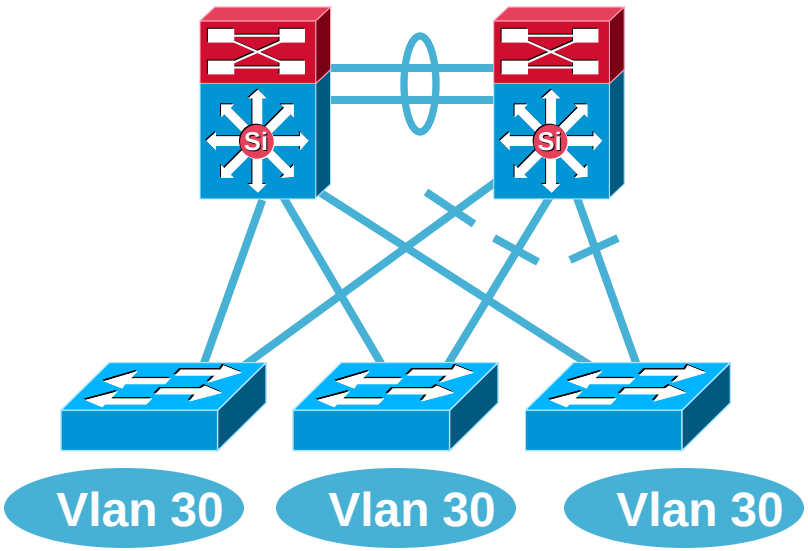
B = STP Blocked Link

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Now We Will Focus on the Unique VLAN per Switch Model



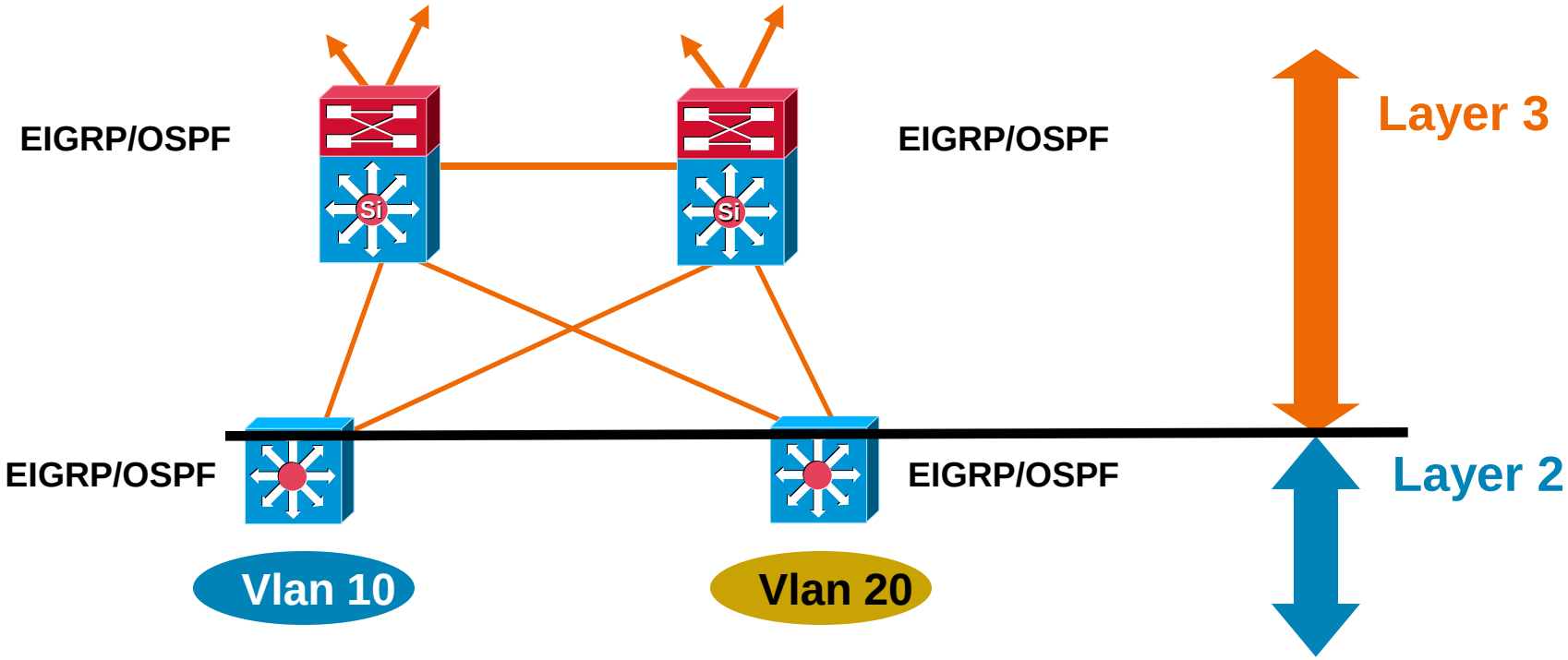
- Each access switch has unique VLAN's
- No layer 2 loops
- Layer 3 link between distribution
- No blocked links



- At least some VLAN's span multiple access switches
- Layer 2 loops
- Layer 2 and 3 running over link between distribution
- Blocked links

© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Confidential

The Best Deployment for This Model Is Routed Access



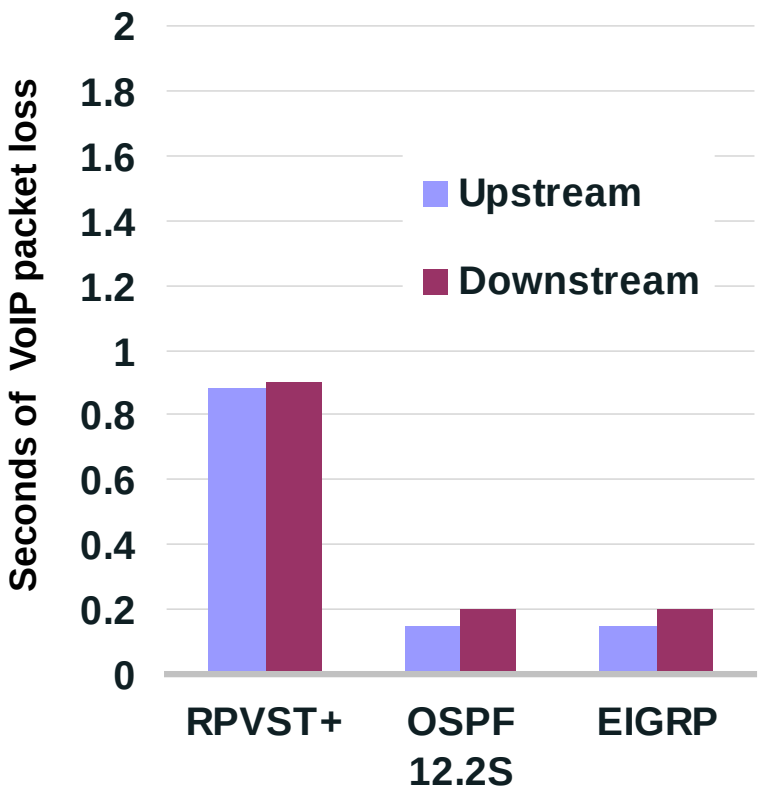
- Extend convergence benefits of ECMP to the edge with Routed Access
- Improved convergence with fewer protocols
- No need to manage FHRP

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Routed Access Provides Rapid Convergence with Optimized Traffic Flow and Ease of Mgmt

- EIGRP converges in <200 msec
- OSPF with sub-second tuning converges in <200 msec

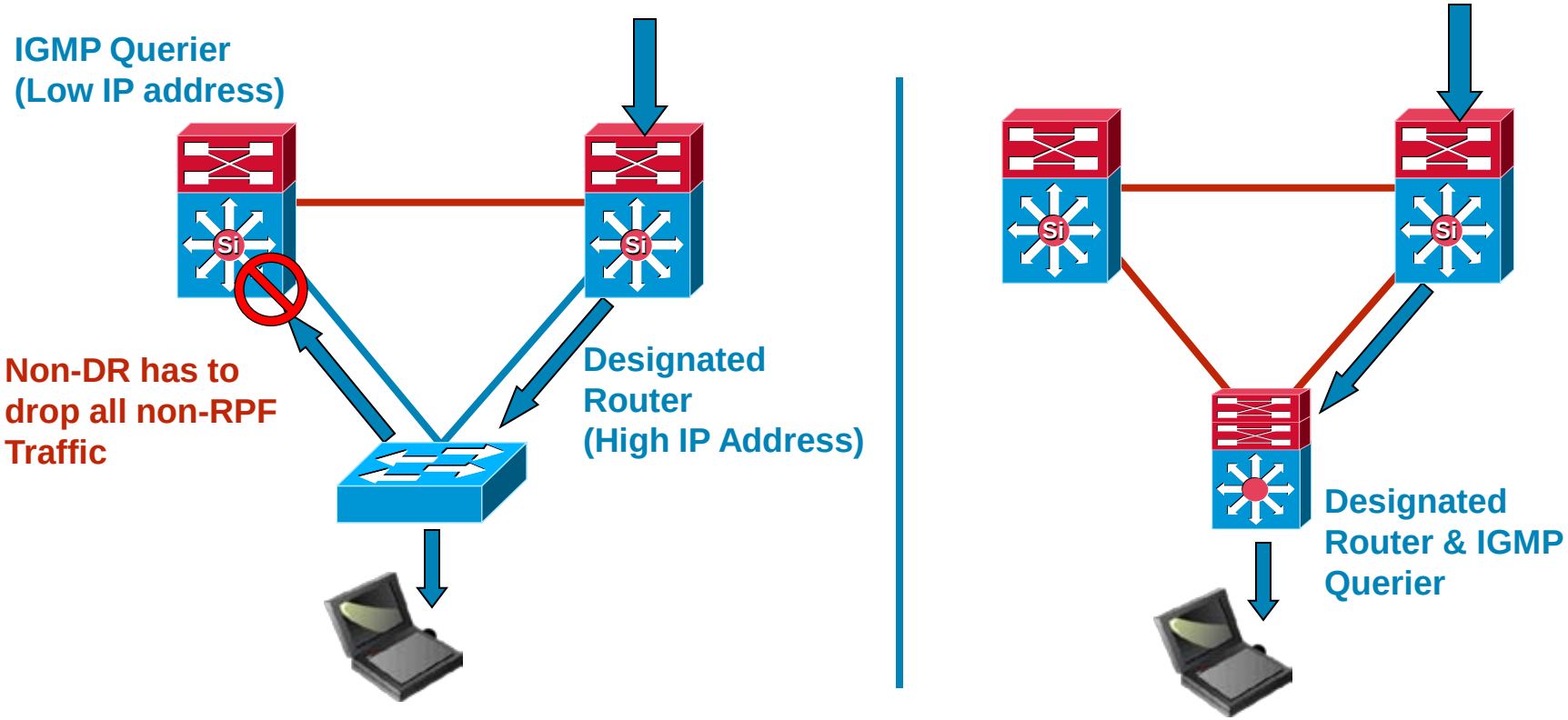
Both L2 and L3 Can Provide Sub-Second Convergence



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

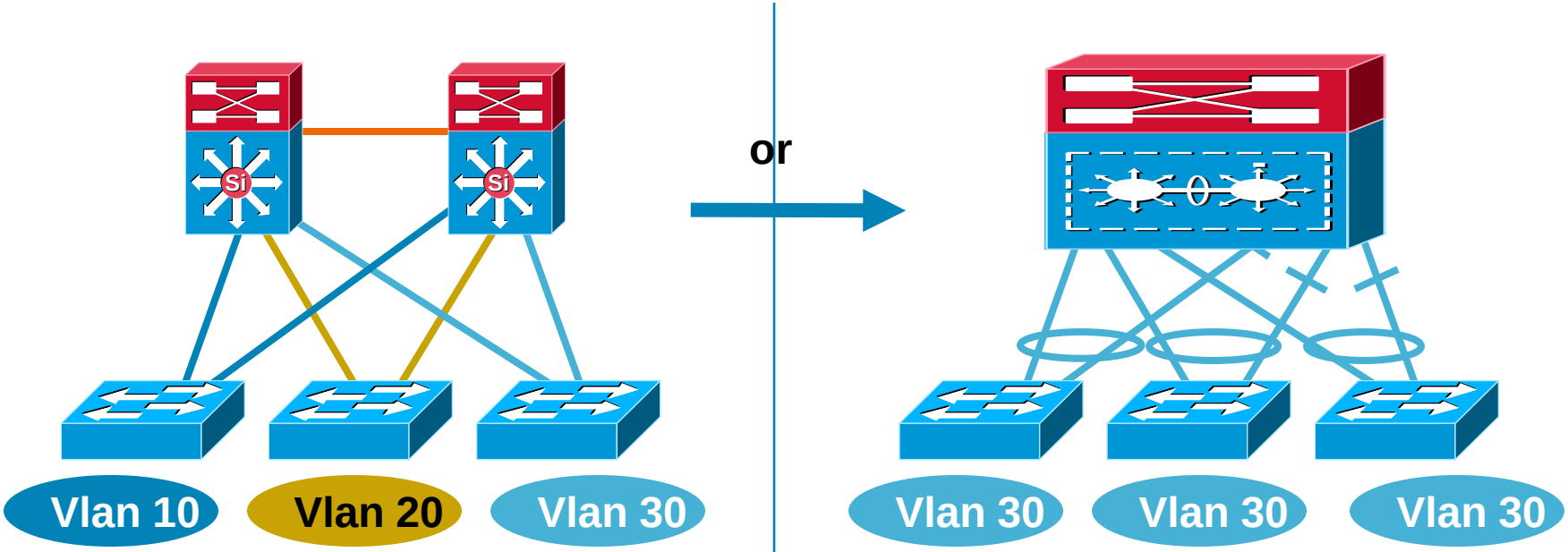
Routed Access Optimized Multicast Operation

- Layer 2 access has two multicast routers on the access subnet, causing one to have to discard frames
- Routed Access has a single multicast router which simplifies management of multicast topology



© 2010 Cisco and/or its affiliates. All rights reserved.

Evolution Possibilities



Use Routed Access

Use MultiChassis Etherchannel with Virtual FHRP

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

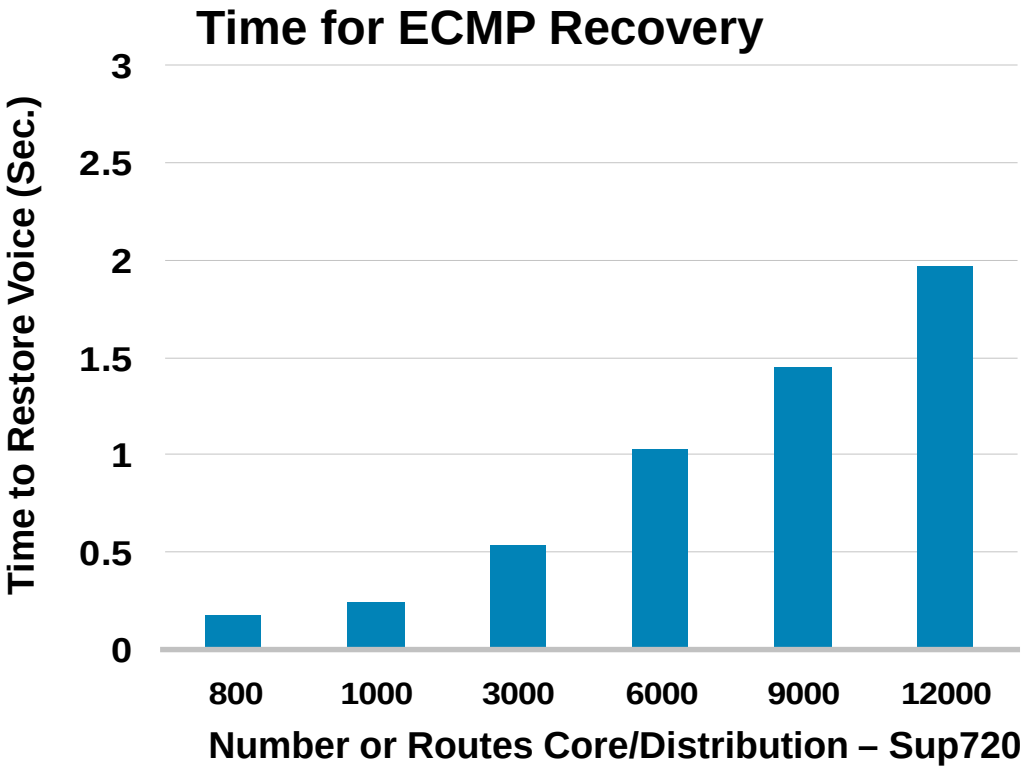
High Availability Campus Design
Agenda

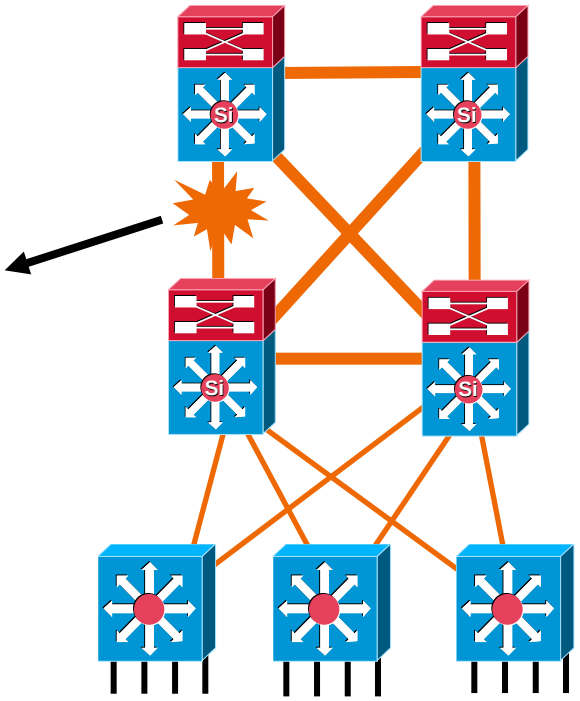
- **Network Level Resiliency**
  - High Availability Design Principles
  - Redundancy in the Distribution Block
  - Campus Routing Best Practices
- **System Level Resiliency**
  - Integrated Hardware and Software Resiliency
    - **NSF/SSO**
    - **ISSU**
  - System Management Resiliency
    - **GOLD & EEM**
- **Hardening the Campus Network Design**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Managing the Number of Routes Directly Affects Convergence Time

- Managing the number of routes in the network is important
- Both EIGRP and OSPF need summarization
- Does not apply to Multi-chassis Etherchannel designs

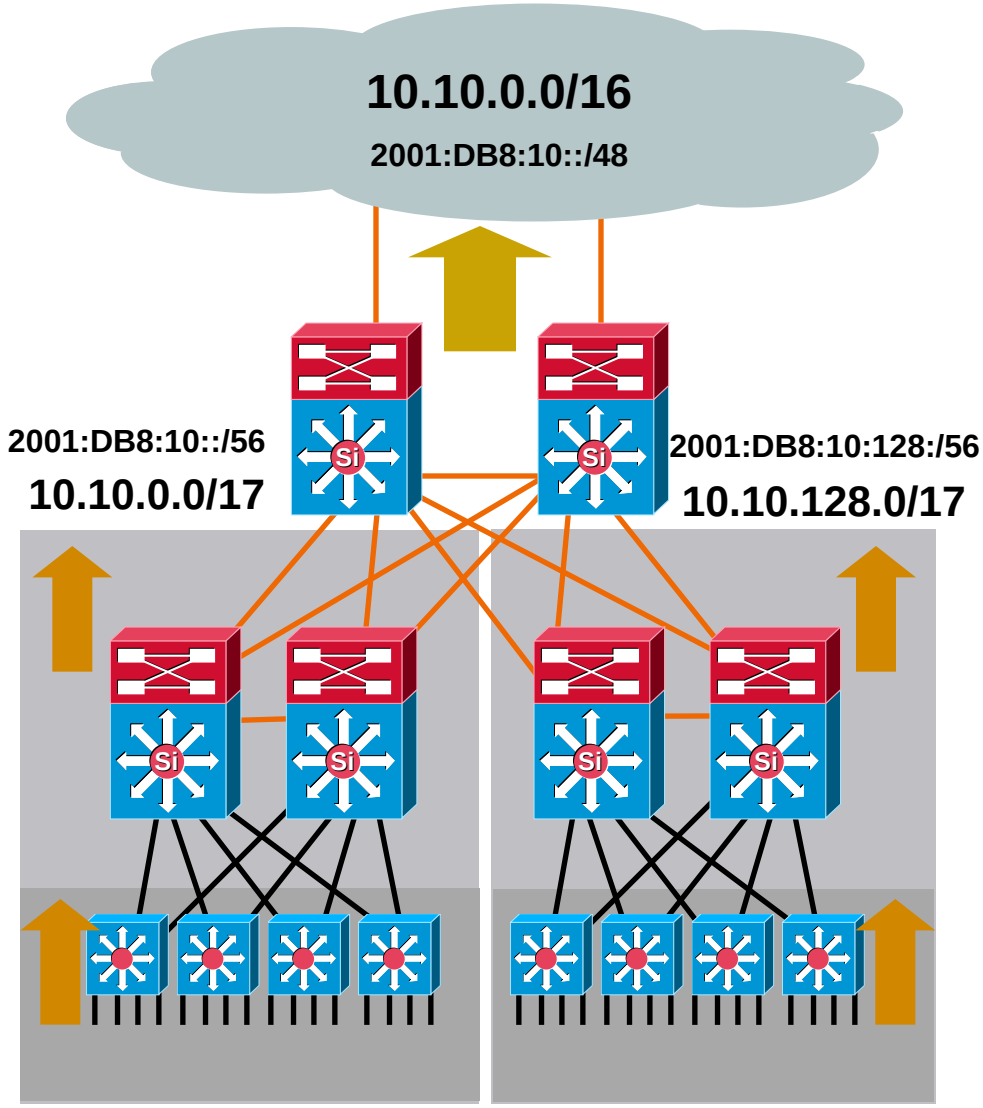




© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

EIGRP Is Unique with It's Multi-Level Summarization Capability

- The greatest advantages of EIGRP are gained when the network has a structured addressing plan that allows for use of summarization and stub routers
- EIGRP provides the ability to implement multiple tiers of summarization and route filtering
- Able to maintain a deterministic convergence time in very large L3 topology



© 2010 Cisco and/or its affiliates. All rights reserved.

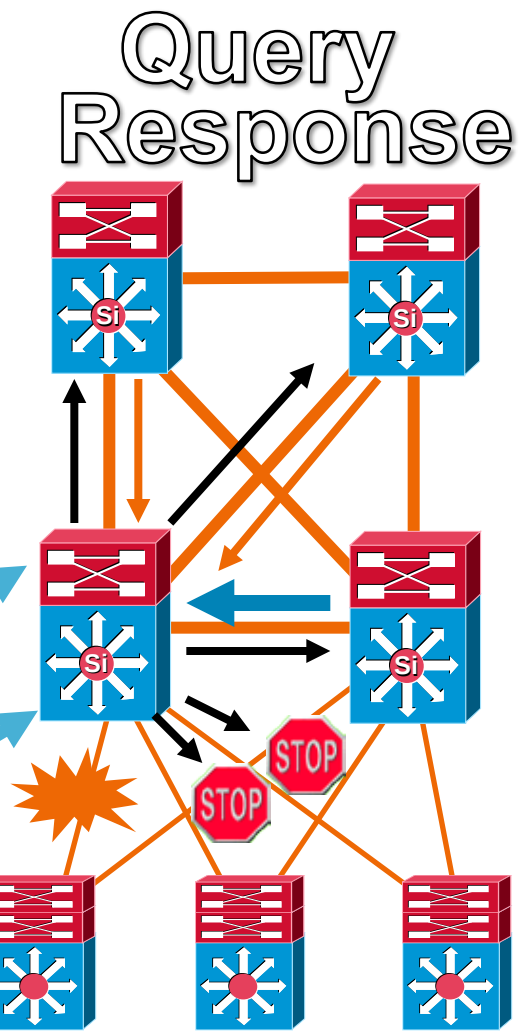
EIGRP Convergence Is Improved with Summarization and Filtering

- EIGRP convergence is largely dependent on query response times
- Minimize the number and time for query response to speed up convergence
- Summarize distribution block routes upstream to the core
- Configure all L3 access switches as EIGRP stub routers
- Filter routes sent down to L3 access switches

```
interface TenGigabitEthernet 4/1
 ipv6 eigrp 100
 ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5
 ipv6 summary-address eigrp 100 2001:DB8:10:120::/56 5
```

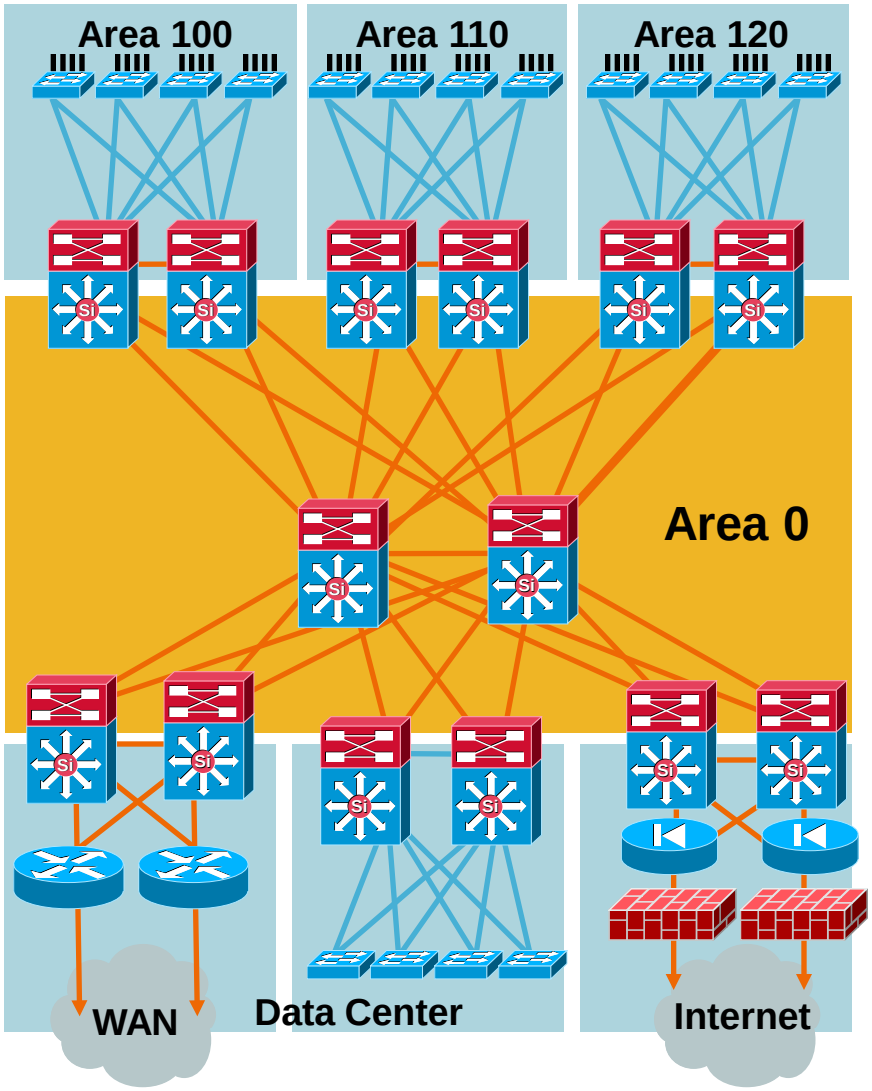
```
router eigrp 100
 network 10.0.0.0
 distribute-list Default out <mod/port>
ip access-list standard Default
 permit 0.0.0.0
!
ipv6 router eigrp 100
 distribute-list prefix-list Default out <mod/port>
ipv6 prefix-list Default permit ::/0
```

```
router eigrp 100
 network 10.0.0.0
 eigrp stub connected
!
ipv6 router eigrp 100
 eigrp stub connected
```



OSPF Area Boundaries Offer Summarization for Improved Scale

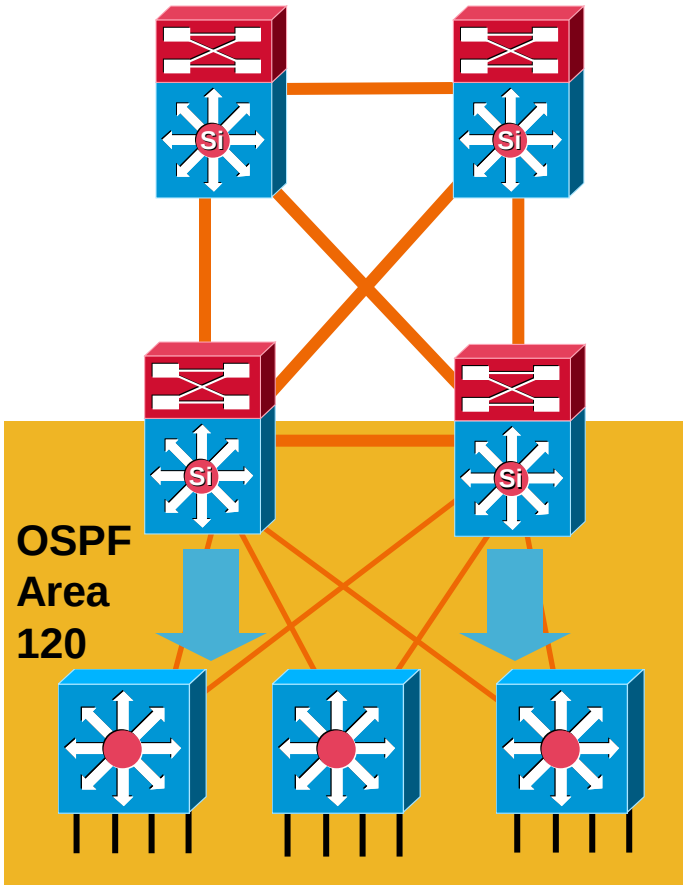
- Area boundaries provide buffers between fault domains
- Keep area 0 for core infrastructure
- Do not extend area 0 to the access routers when using Routed Access



© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Confidential

OSPF Downstream Summarization Is Accomplished with Multiple Area Types

- ABR for a regular area forwards
  - **Summary LSAs (Type 3)**
  - **ASBR summary (Type 4)**
  - **Specific externals (Type 5)**
- Stub area ABR forwards
  - **Summary LSAs (Type 3)**
  - **Summary default (0.0.0.0 - ::/0)**
- A totally stubby area ABR forwards
  - **Summary default (0.0.0.0 - ::/0)**




```
router ospf 100
 area 120 stub no-summary
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

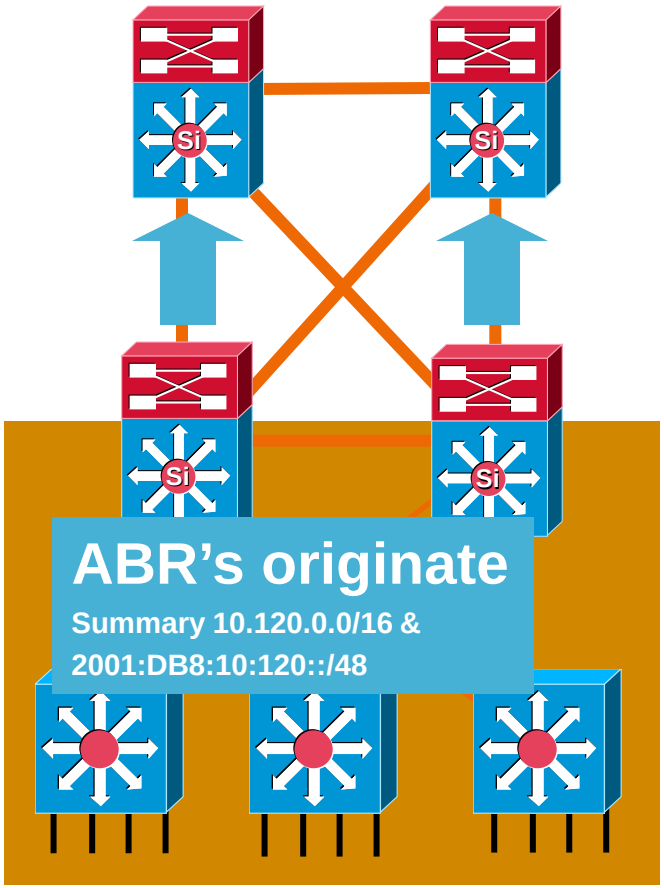
```
ipv6 router ospf 100
area 120 stub no-summary
!
interface Vlan120
 ipv6 address 2001:DB8:10:120::/64 eui-64
 ipv6 ospf 100 area 120
```

OSPF Upstream Summarization Helps Minimize LSA Churn in the Core

- Summarize routes from the distribution block upstream into the core
- Minimize the number of LSA's and routes in the core
- Reduce the need for SPF calculations due to internal distribution block changes

```
router ospf 100
 area 120 stub no-summary
 area 120 range 10.120.0.0 255.255.0.0 cost 10
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

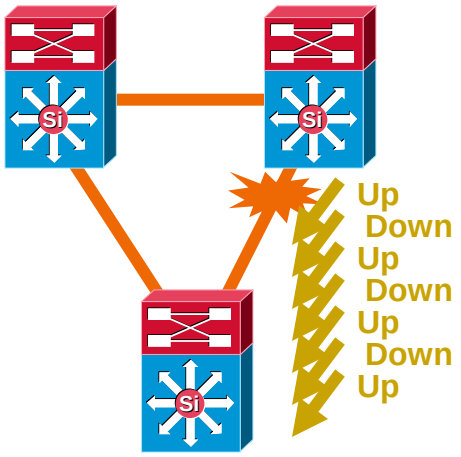
```
interface Vlan120
 ipv6 address 2001:DB8:10:120::/64 eui-64
 ipv6 ospf 100 area 120
!
ipv6 router ospf 100
 area 120 stub no-summary
 area 120 range 2001:DB8:10:120::/56 cost 10
```



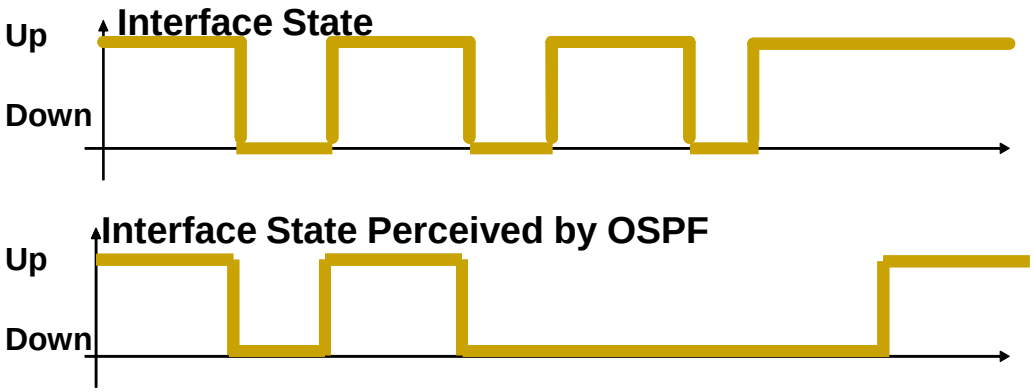
© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Routing Protocol Churn Can Be Reduced with IP Event Dampening

- Prevents routing protocol churn caused by constant interface state changes
- Dampening is applied on a system: nothing is exchanged between routing protocols
- Supports all IP routing protocols
  - Static routing, RIP, EIGRP, OSPF, IS-IS, BGP
  - In addition, it supports HSRP and CLNS routing
  - Applies on physical interfaces and can't be applied on subinterfaces individually




```
interface GigabitEthernet1/1
 description Uplink to Distribution 1
 dampening
 ip address 10.120.0.205 255.255.255.254
 ipv6 address 2001:DB8:10:120::205/64 eui-64
```



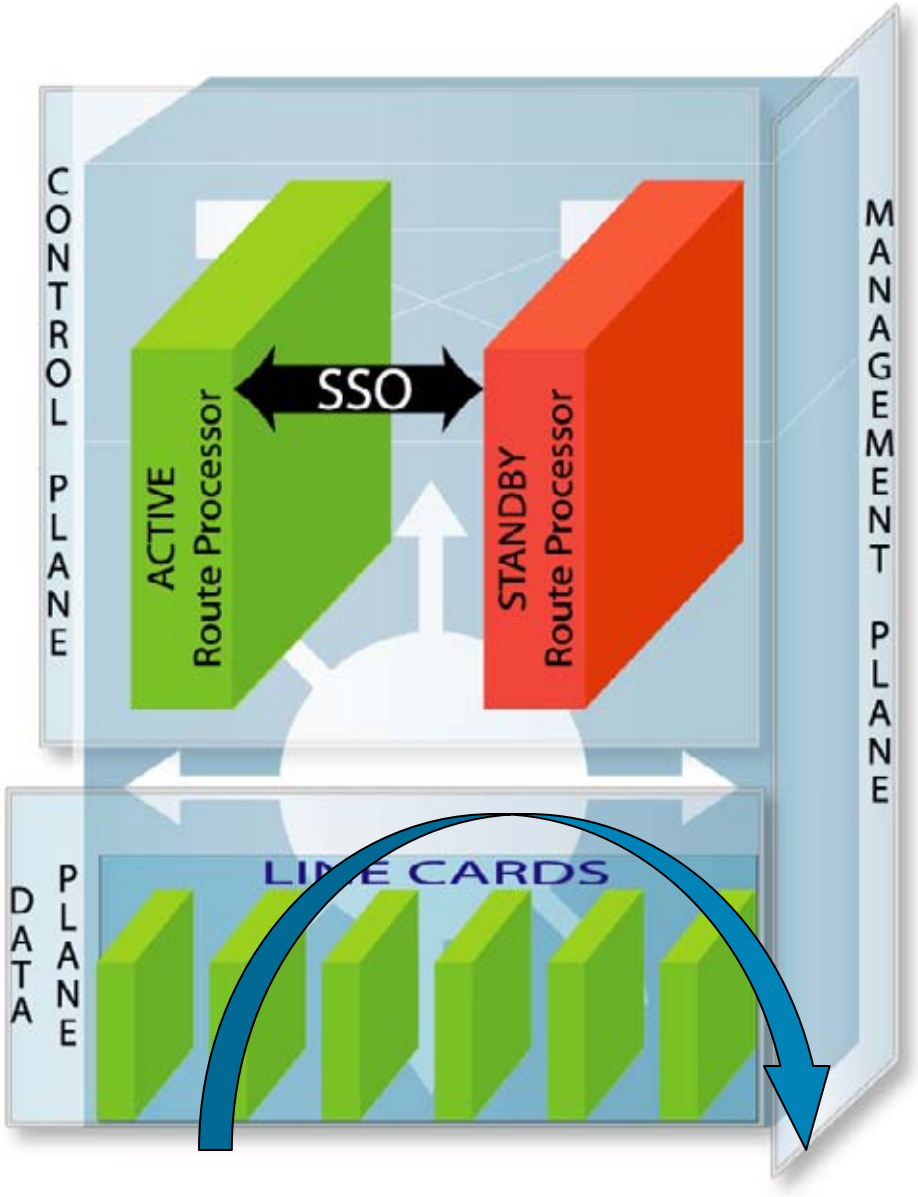
© 2010 Cisco and/or its affiliates. All rights reserved.

High Availability Campus Design
Agenda

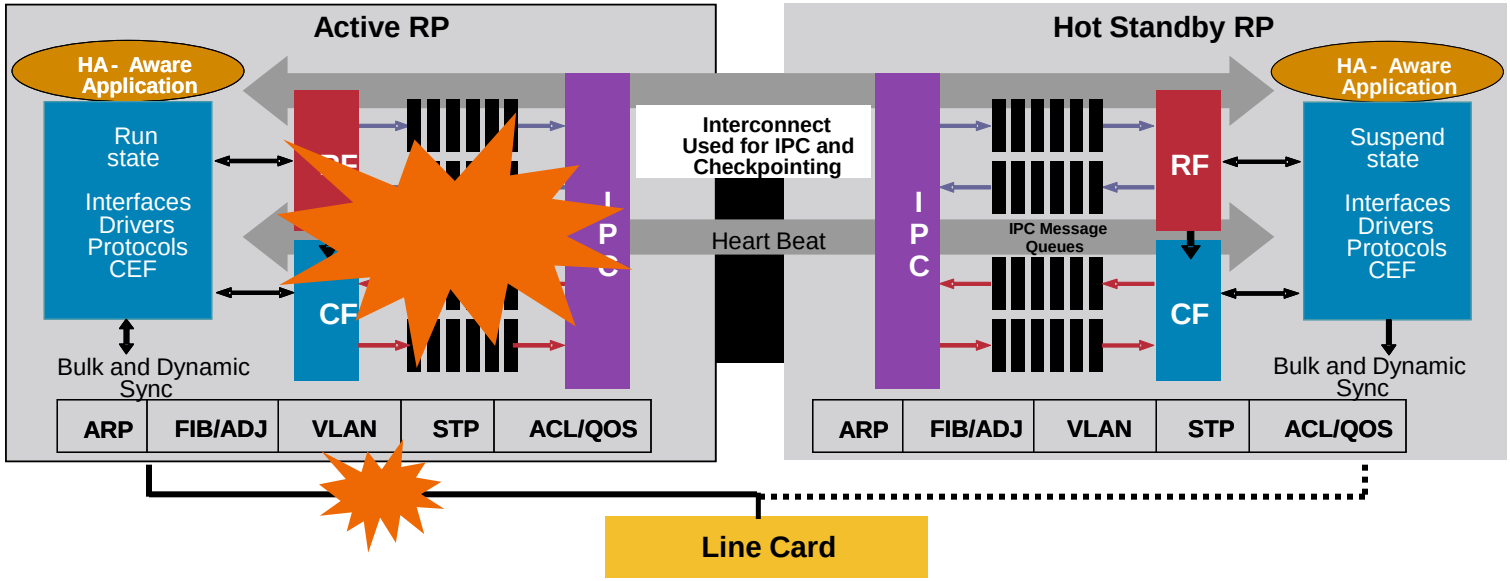
- **Network Level Resiliency**
  - High Availability Design Principles
  - Redundancy in the Distribution Block
  - Campus Routing Best Practices
- **System Level Resiliency**
  - Integrated Hardware and Software Resiliency
    - **NSF/SSO**
    - **ISSU**
  - System Management Resiliency
    - **GOLD & EEM**
- **Hardening the Campus Network Design**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Supervisor Redundancy Is Provided by Stateful Switch Over



Stateful Switchover (SSO) Keeps Both Processors in Synch for Hot Standby Mode



- Active/Standby supervisors run in synchronized mode (boot-env, running-configuration, protocol state and line cards status gets synchronized)
- Depending on platform, line card and protocol this incurs from 0 to 3 seconds of outage
- Switch processors synchronize L2 information, (e.g., STP, MAC address) and L2/L3 FIB, QOS and ACL tables
- DFCs are populated with L2/L3 FIB, Netflow, and ACL tables
- Line card protocol status is maintained during failover

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Stackwise+ Provides Stack Master Switch Redundancy

- 1:N redundancy

  Any member can become stack master if the master fails. The switch with the highest stack-member priority-value becomes the master using RPR+ and NSF

- Stack master provides centralized functionality

  Layer 3 control plane, management plane (telnet, SSH, SNMP, HTTP), and stack configuration

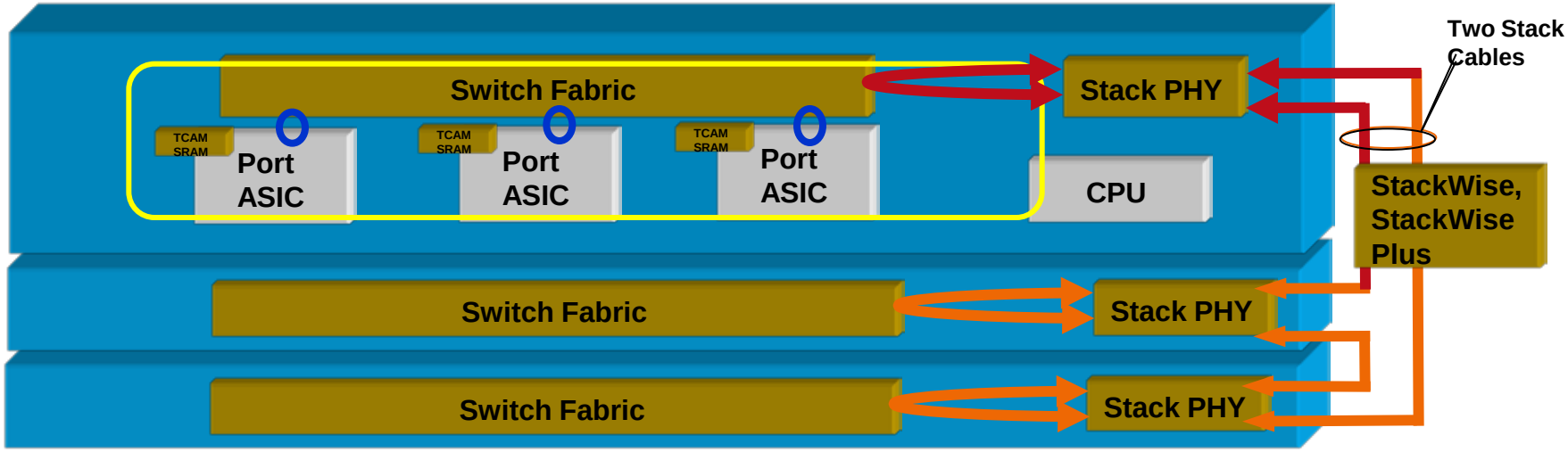
  Controls CDP and EtherChannel and propagates VLAN database and Spanning Tree info

  Builds and propagates the hardware information (L3 FIB, ACL, QoS)

  Control routing process and neighbor adjacency for L3 routed operation

- Stack member provides distributed forwarding of data and local control plane

  Local instance of STP, BPDU processing, and MAC address management



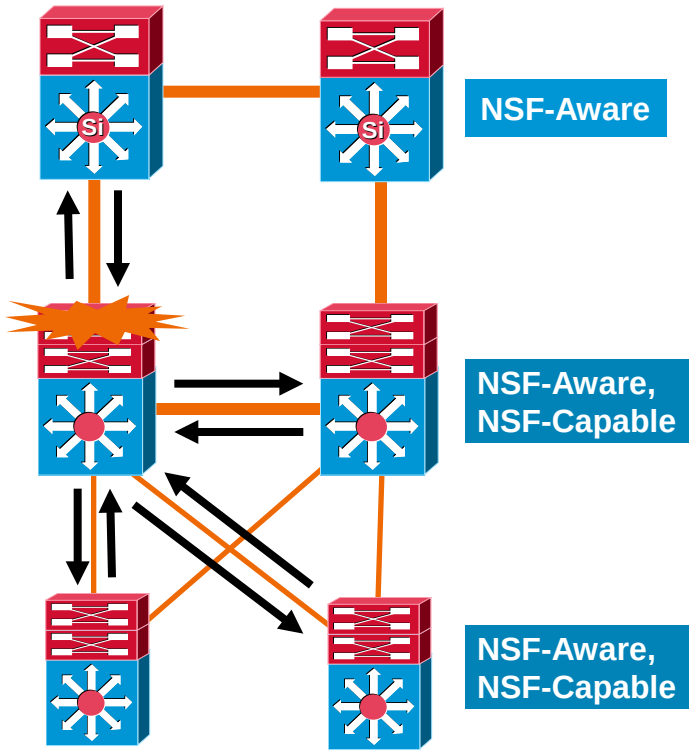
© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

NSF Works with SSO to Keep Neighbors Forwarding During a Supervisor Switchover

- Non-Stop Forwarding provides graceful restart enhancements to EIGRP, OSPF, IS-IS and BGP
- An NSF-capable router continuously forwards packets during an SSO processor recovery
- NSF-aware and NSF-capable routers provide for transparent routing protocol recovery

  Graceful restart extensions enable neighbor recovery without resetting adjacencies

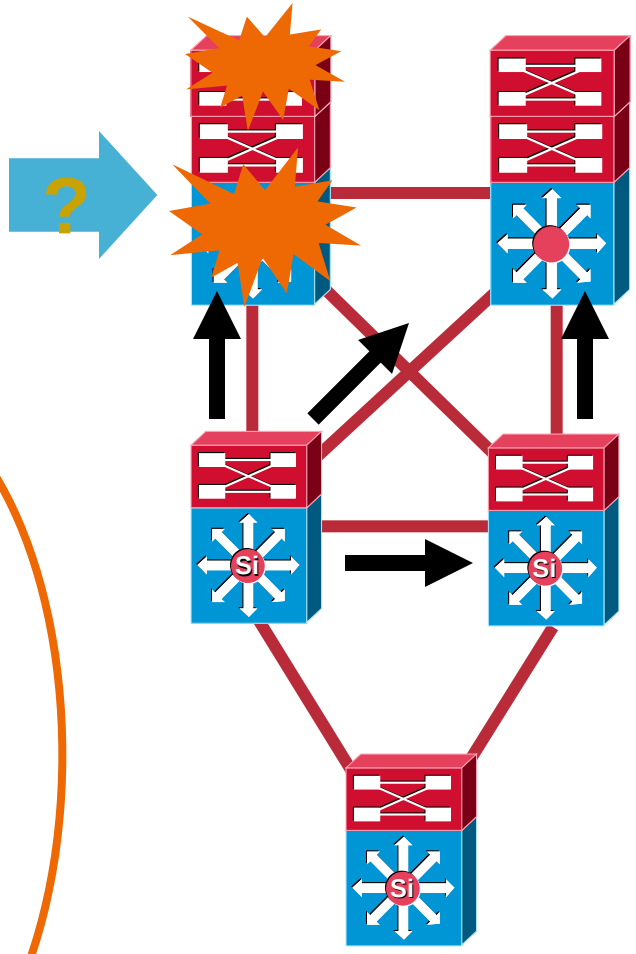
  Routing database re-synchronization occurs in the background

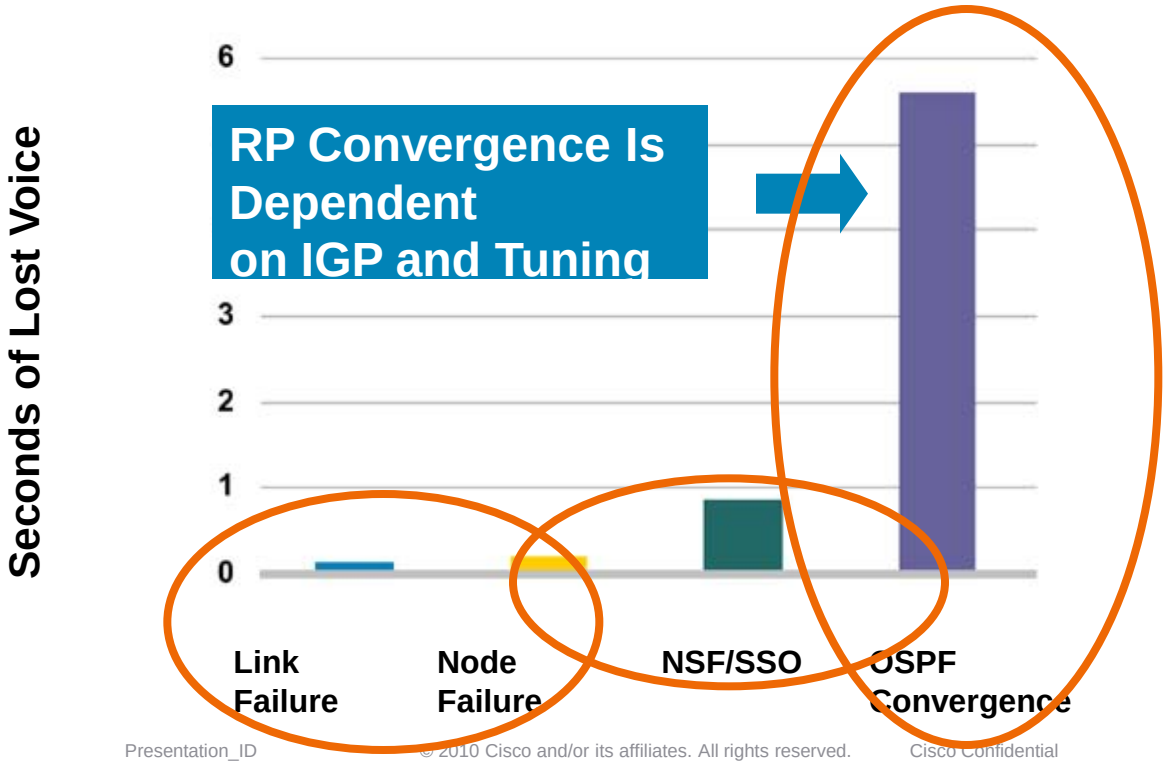


© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

NSF/SSO Offers Better Convergence in Non Full Mesh Topologies

- Redundant topologies with equal cost paths provide sub-second convergence
- NSF/SSO provides superior availability in environments with non-redundant paths





© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

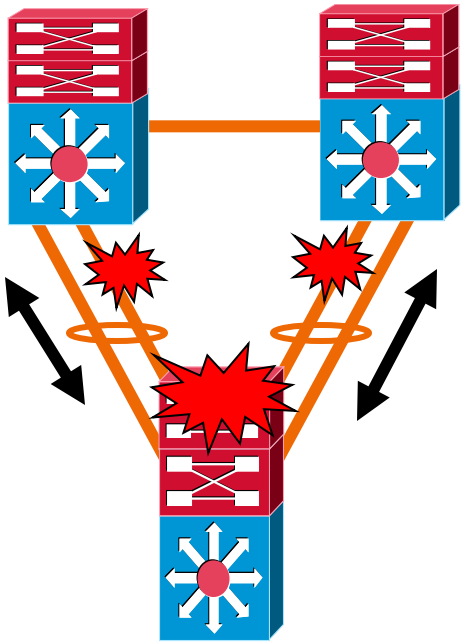
If Using the Supervisor Uplinks, EtherChannel Is Recommended

- The use of uplinks with NSF/SSO results in a dual component failure

  Supervisor Failure + Supervisor Ports Failure

- To avoid network protocol convergence use EtherChannel across supervisor uplinks or simply use line card ports as uplinks

No Traffic Flow Change

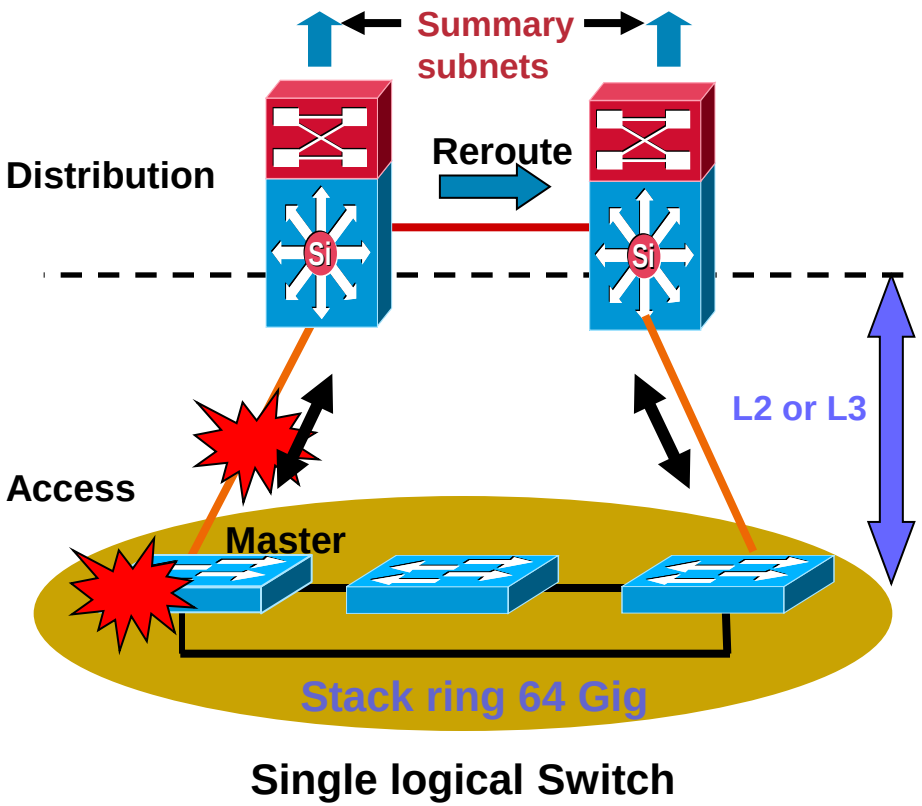


EtherChannel or Uplink from line card

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Avoid Using Uplinks on Stack Master in a Stackwise + Design

- Using uplink ports on non-Master switches avoids unnecessary network protocol convergence when a master fails
- You want to basically avoid having 2 failures when the master fails
- Convergence impact depends on tuning but could be up to 7 seconds.
- Recommendation is *NOT* to combine Master and uplink in single switch
- Use Multi-chassis etherchannel to minimize this failure



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

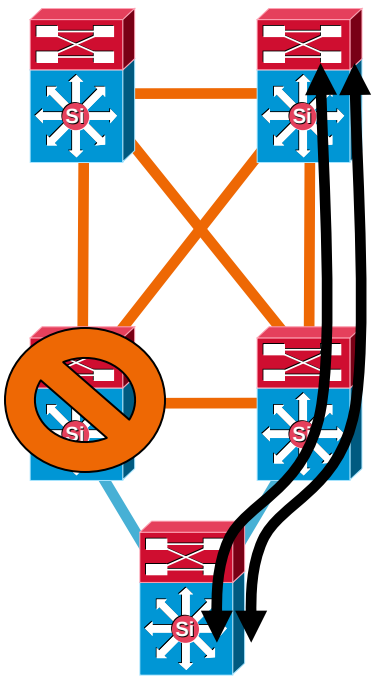
High Availability Campus Design
Agenda

- **Network Level Resiliency**
  - High Availability Design Principles
  - Redundancy in the Distribution Block
  - Campus Routing Best Practices
- **System Level Resiliency**
  - Integrated Hardware and Software Resiliency
    - **NSF/SSO**
    - **ISSU**
  - System Management Resiliency
    - **GOLD & EEM**
- **Hardening the Campus Network Design**

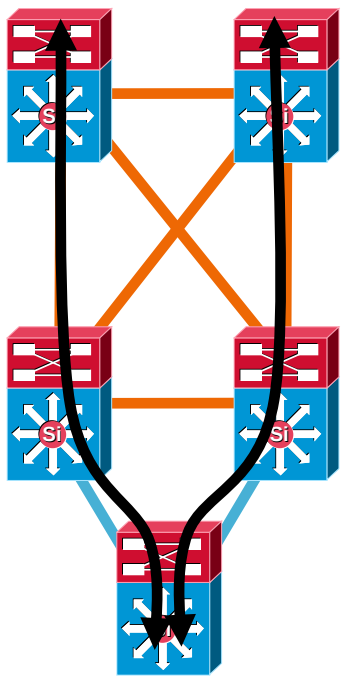
© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

In Service Software Upgrade Allows Upgrade Without Taking Switch Down

- In redundant topology standard maintenance practice is to shut down devices during upgrade and let the network converge
- ISSU provides the ability to upgrade software in place without having to shut down
- Offers significant uptime improvements



Scheduled Maintenance—Half Capacity

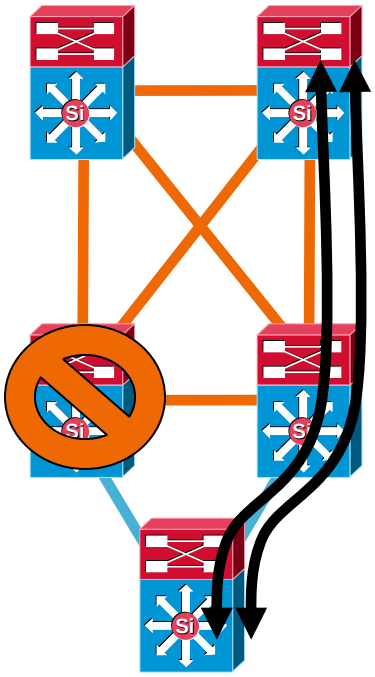


ISSU—All Paths and Switches Active During Upgrade

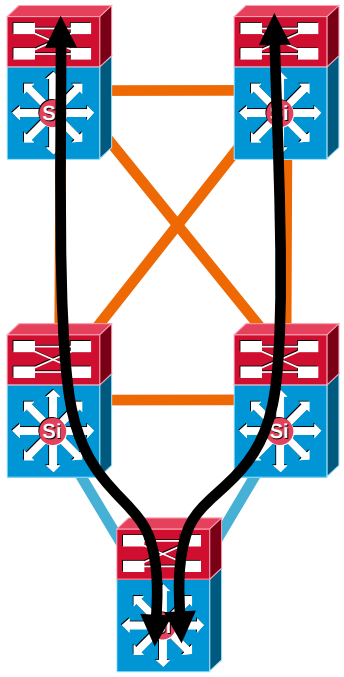
© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

System Image Software Upgrade Downtime Varies By Chassis

- Catalyst 6500 software upgrade with enhanced Fast Software Upgrade enabled line cards to load software with 60+ seconds of downtime
- Catalyst 4500 with In-Service Software Upgrade will load software with 10 to 30 ms of downtime
- Nexus 7000 with In-Service Software Upgrade will load software with no downtime



Scheduled Maintenance—Half Capacity



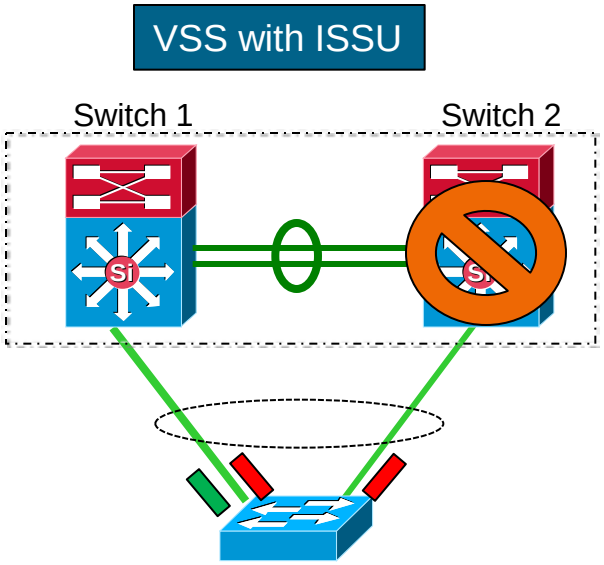
ISSU—All Paths and Switches Active During Upgrade

© 2010 Cisco and/or its affiliates. All rights reserved.

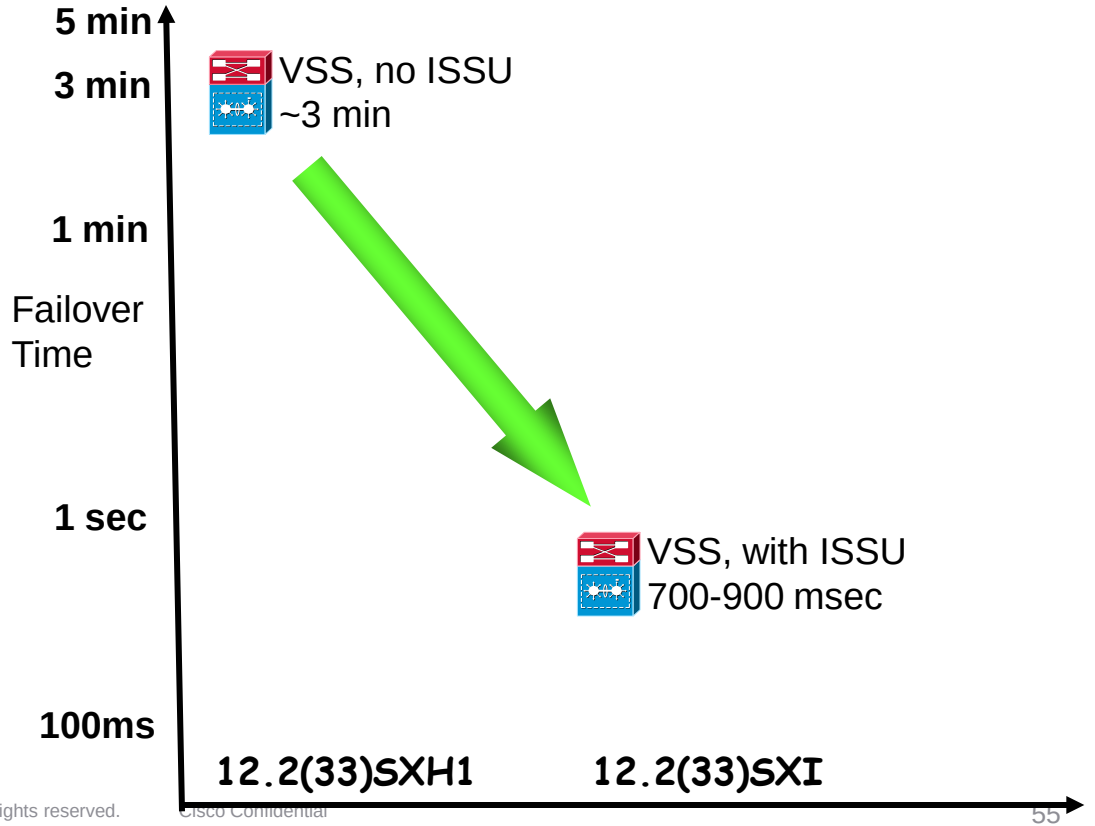
Multichassis EtherChannel Automates Half Capacity in ~300 msec

Cisco ISSU Phase 1 (In-Service-Software-Upgrade) & VSS 12.2(33)SXI

- Greatly reduces SW upgrade maintenance window (however still requires linecard reload)
- VSS & ISSU will provide ~1000msec upgrade window (assuming dual-homed access devices)
- Very suitable for voice deployments, where acceptable loss of service is close to 300 msec



Total network impact – About 1 second



© 2010 Cisco and/or its affiliates. All rights reserved.
Cisco Confidential

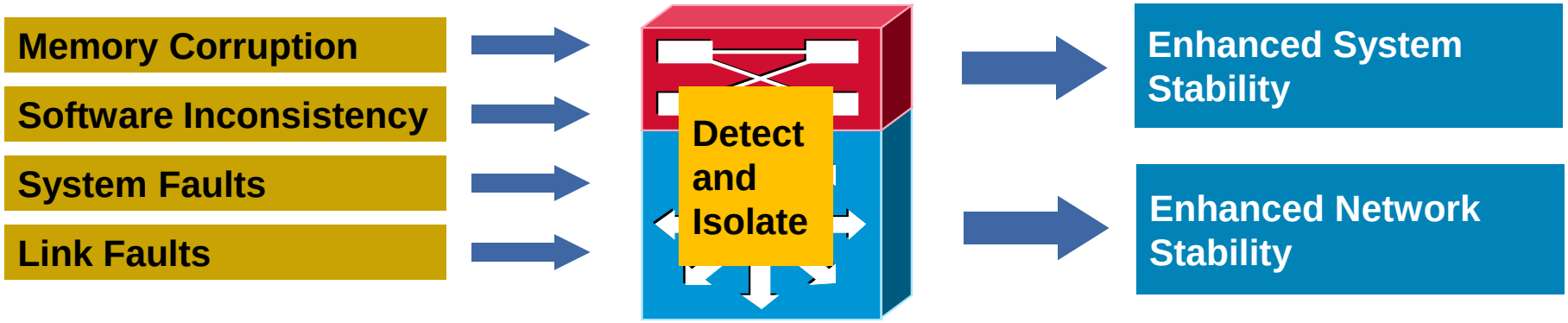
High Availability Campus Design
Agenda

- **Network Level Resiliency**
  - High Availability Design Principles
  - Redundancy in the Distribution Block
  - Campus Routing Best Practices
- **System Level Resiliency**
  - Integrated Hardware and Software Resiliency
    - **NSF/SSO**
    - **ISSU**
  - System Management Resiliency
    - **GOLD & EEM**
- **Hardening the Campus Network Design**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

GOLD and EEM Offer Proactive Fault Detection and Reaction

- Challenge: In today's highly available networks improved physical redundancy is not enough, intelligent system failure detection and recovery are key



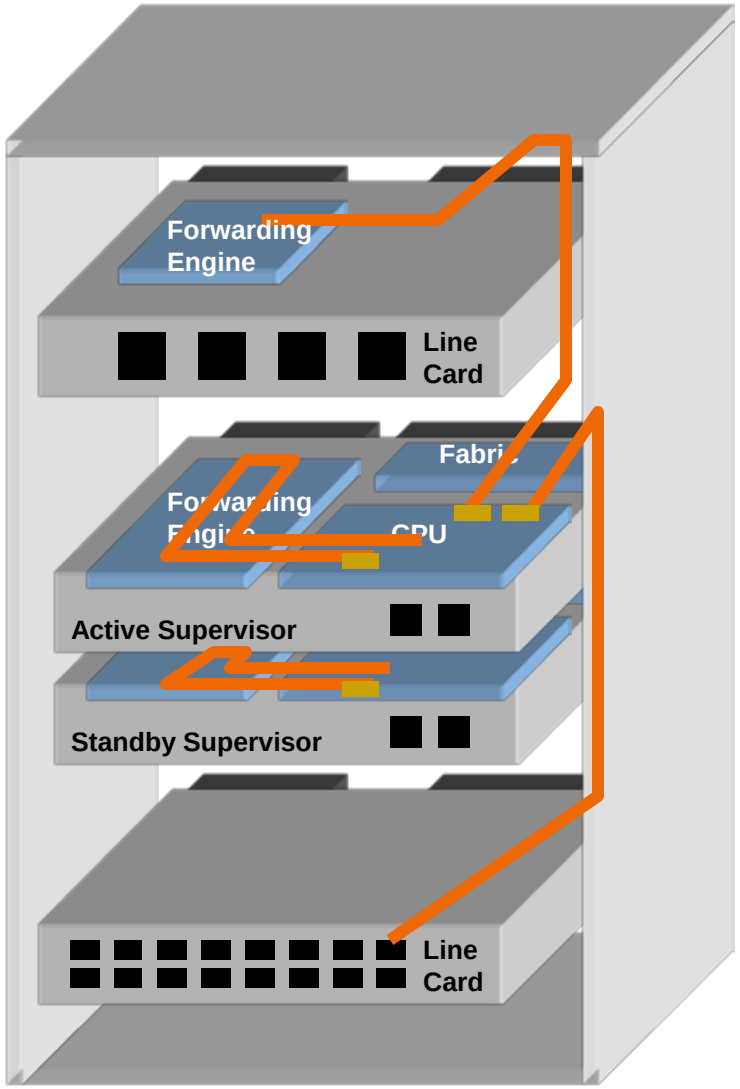
Generic Online Diagnostics (GOLD) Provides Proactive, Scheduled and Manual System Diagnostics

Enhanced Object Tracking (EOT), Embedded Event Manager (EEM), and Smart Call Home (SCH) Provide Intelligent Response to System Events

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Generic Online Diagnostics Verifies Operation of Data Plane and Control Plane

- GOLD: Common Framework to check the health of hardware components and verify proper operation of the system data plane and control plane at run-time and boot-time
- Diagnostic packet switching tests verify that the system is operating correctly:
  - Is the supervisor control plane and forwarding plane functioning properly?
  - Is the standby supervisor ready to take over?
  - Are linecards forwarding packets properly?
  - Are all ports working?
  - Is the backplane connection working?
- Other types of diagnostics tests including memory and error correlation tests are also available



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Multiple Tests Are Provided for Different Scenarios

Boot-Up Diagnostics

```
Switch(config)#diagnostic bootup level complete
```

Run During System Bootup, Line Card OIR or Supervisor Switchover
Makes Sure Faulty Hardware Is Taken out of Service

Runtime Diagnostics

Health-Monitoring

```
Switch(config)#diagnostic monitor module 5 test 2
Switch(config)#diagnostic monitor interval module 5 test 2
00:00:15
```

Non-Disruptive Tests Run in the Background
Serves as HA Trigger

On-Demand

```
Switch#diagnostic start module 4 test 8
Module 4: Running test(s) 8 may disrupt normal
system operation
Do you want to continue? [no]: y
Switch#diagnostic stop module 4
```

All Diagnostics Tests Can Be Run on Demand, for Troubleshooting Purposes. It Can Also Be Used As A Pre-deployment Tool

Scheduled

```
Switch(config)#diagnostic schedule module 4
test 1 port 3 on Jan 3 2005 23:32
Switch(config)#diagnostic schedule module 4
test 2 daily 14:45
```

Schedule Diagnostics Tests, for Verification and Troubleshooting Purposes

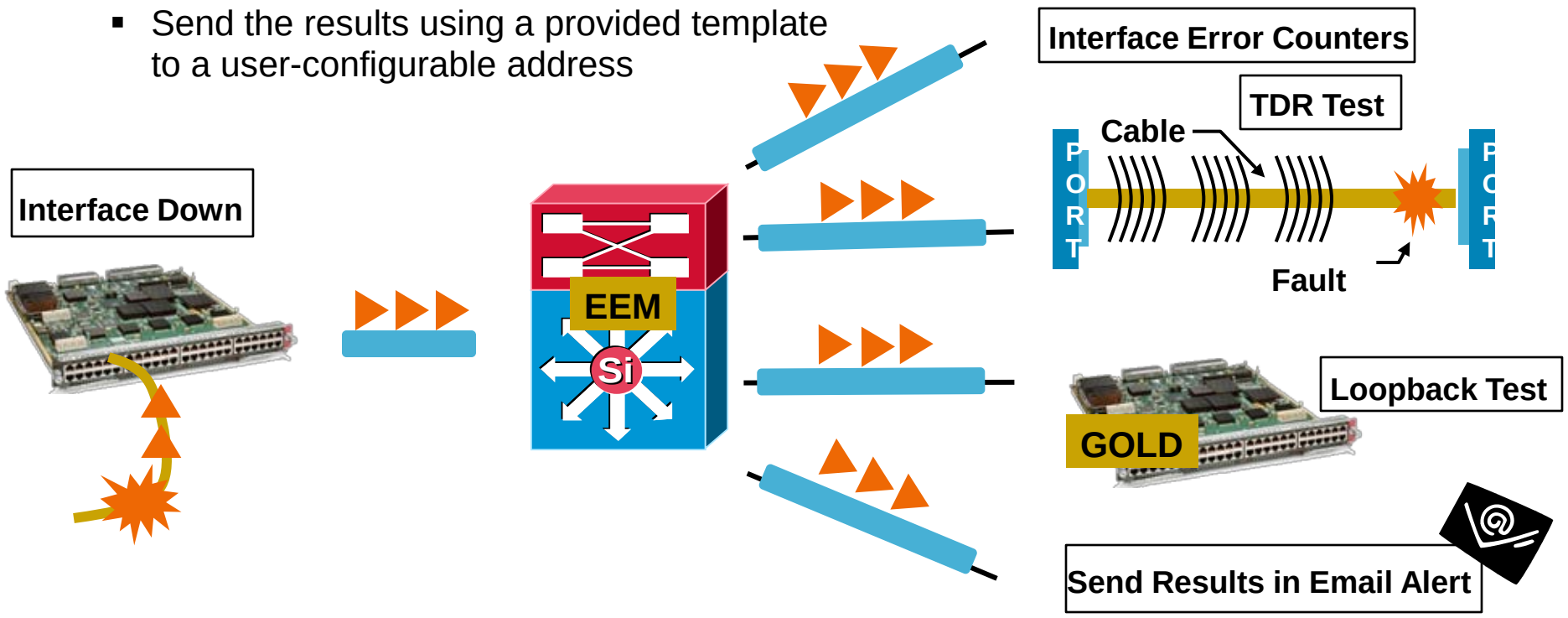
© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Embedded Event Manager

EEM Application Example

Upon Matching the Provided SYSLOG Message 'LINK-3-UPDOWN', the Switch Performs the Following Actions:

- Display error statistics for the link that has gone down
- Start a Time Domain Reflectometry (TDR) test
- Start a GOLD Loopback test
- Send the results using a provided template to a user-configurable address

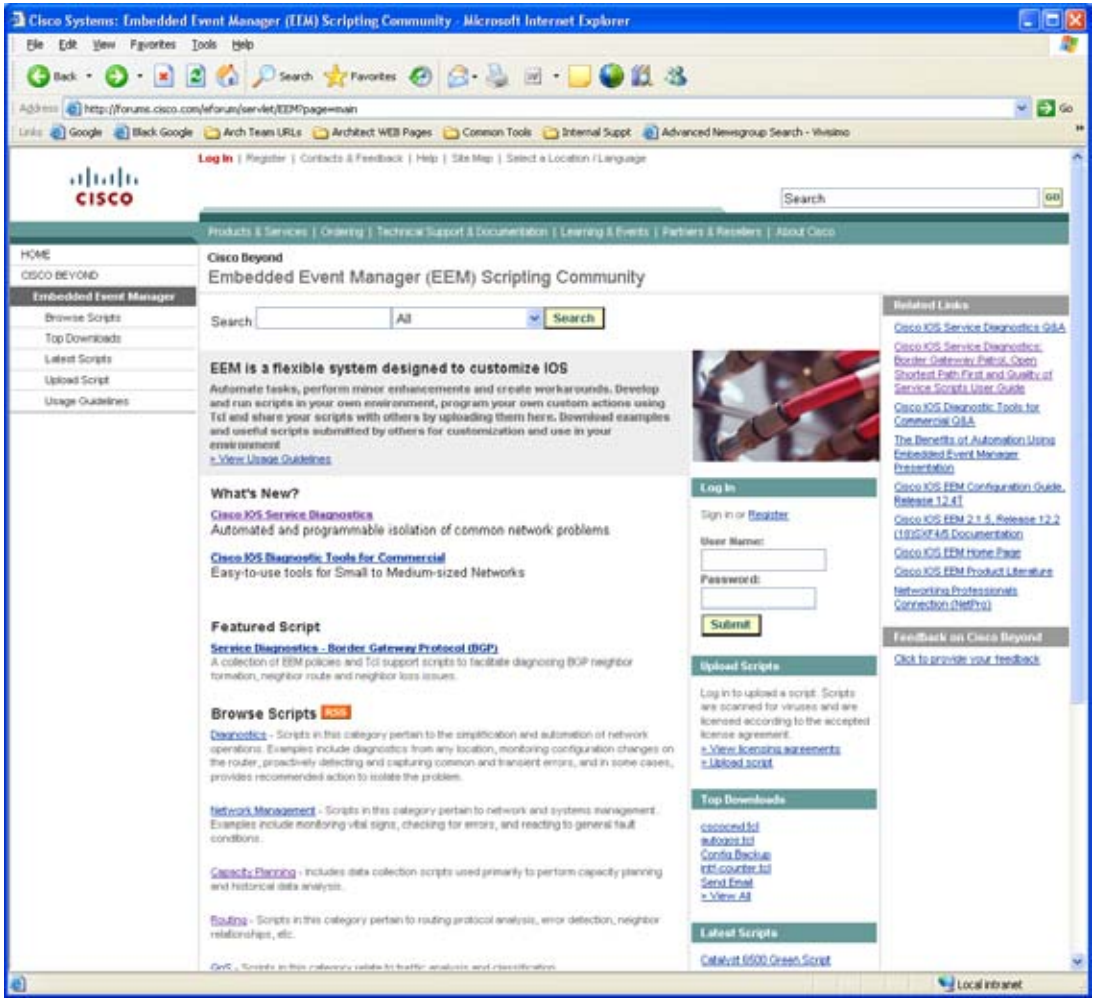


© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Embedded Event Manager

Cisco Beyond: Product Extension Community

- On Cisco.com
- /go/ciscobeyond
- Open source scripts, share, upload, download, learn by example
- Categories include: Ntwk mgmt, routing, QoS, High availability, User interface, etc.
- Scripts can be programmed in IOS CLI or TCL


Cisco Systems: Embedded Event Manager (EEM) Scripting Community - Microsoft Internet Explorer
Embedded Event Manager (EEM) Scripting Community
EEM is a flexible system designed to customize IOS
What's New?
Featured Script
Browse Scripts

http://cisco.com/go/ciscobeyond

© 2010 Cisco and/or its affiliates. All rights reserved.

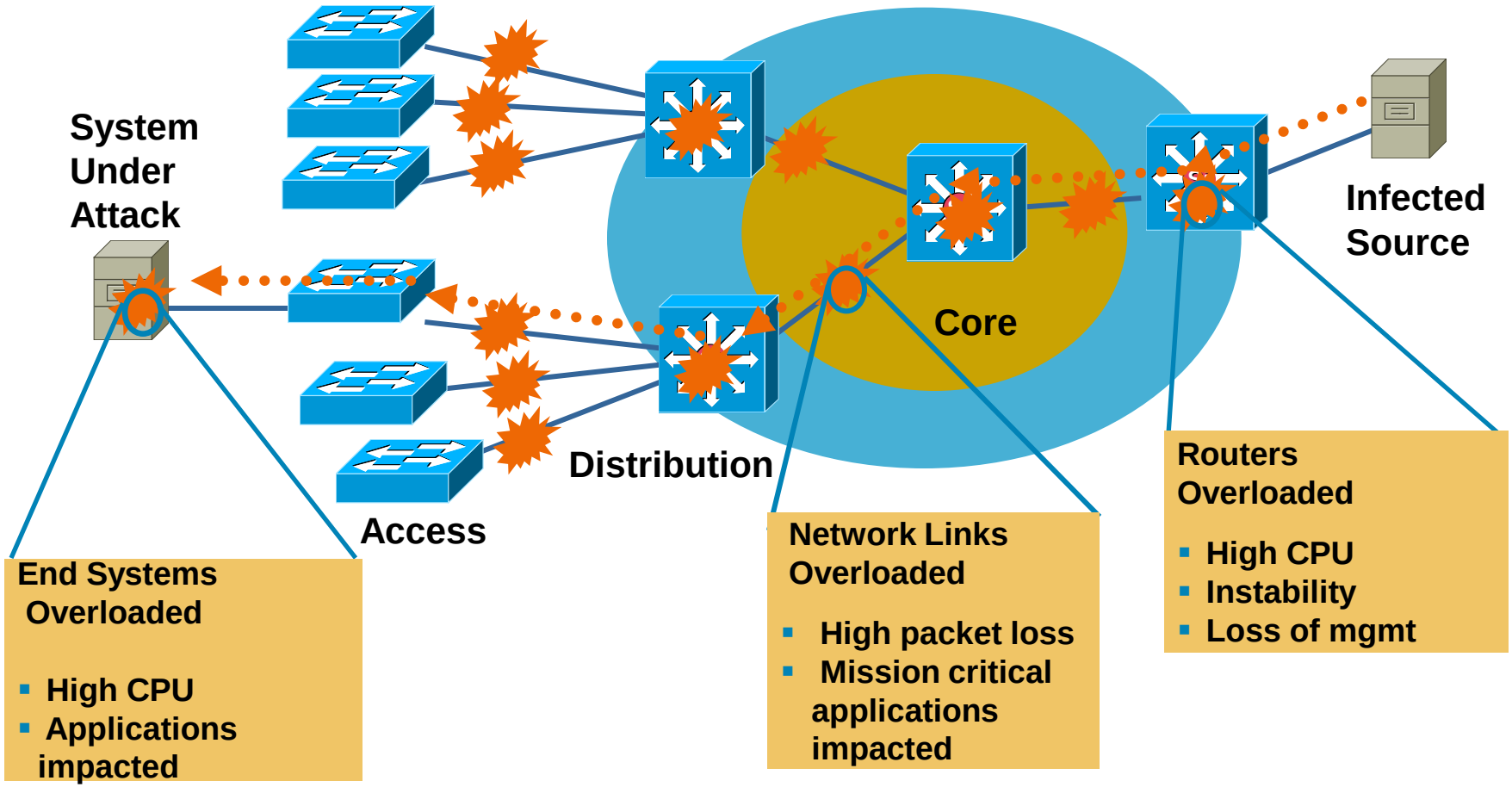
High Availability Campus Design
Agenda

- **Network Level Resiliency**
  - High Availability Design Principles
  - Redundancy in the Distribution Block
  - Campus Routing Best Practices
- **System Level Resiliency**
  - Integrated Hardware and Software Resiliency
    - **NSF/SSO**
    - **ISSU**
  - System Management Resiliency
    - **GOLD & EEM**
- **Hardening the Campus Network Design**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Impact of a Network Attack

Direct and Collateral Damage

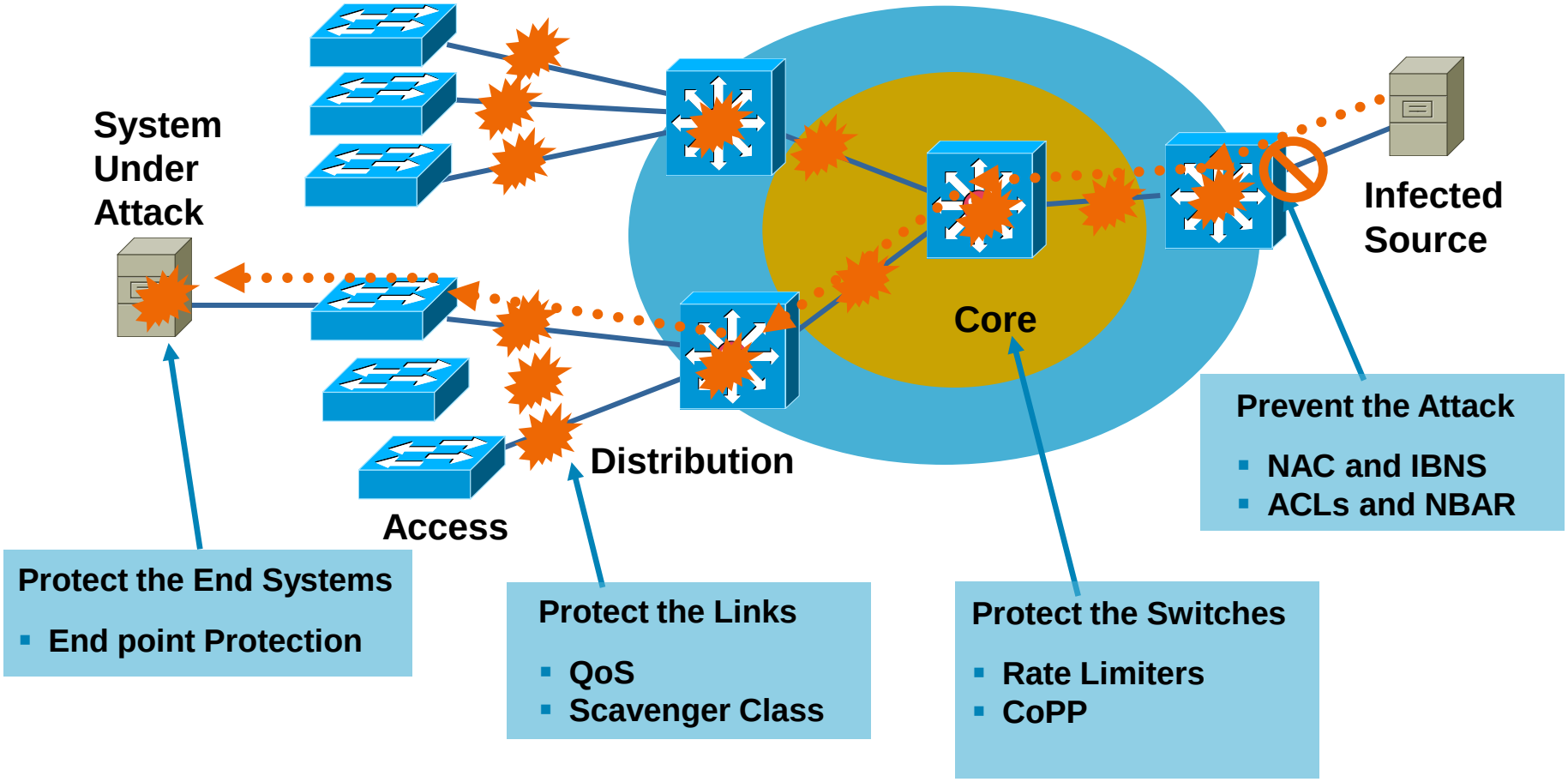


Availability of Networking Resources Impacted by Malicious Activity

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Mitigating Network Attacks

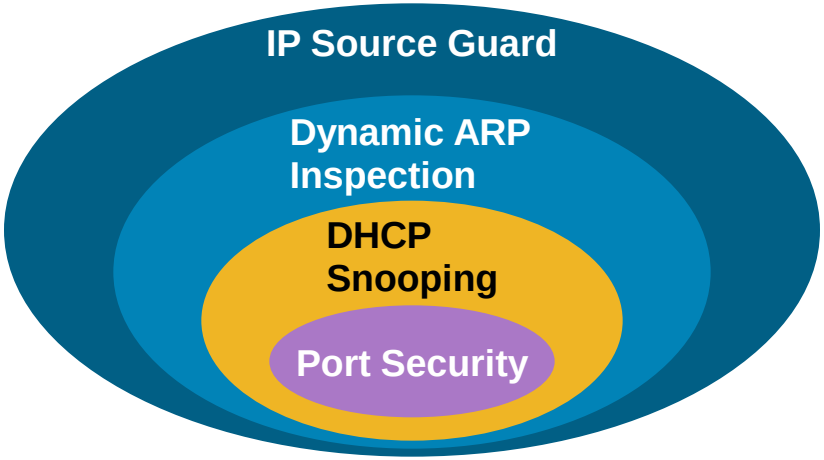
Preventing and Limiting the Pain

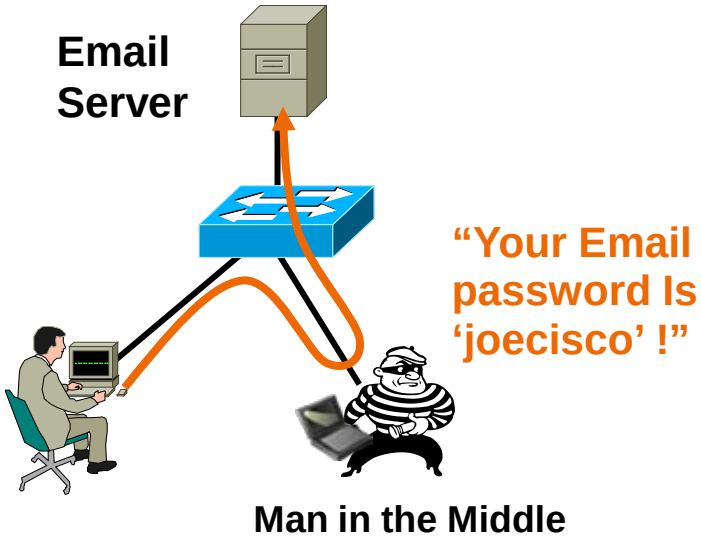


Allow the network to protect itself

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Use a Multilayer Approach to Hardening The Edge



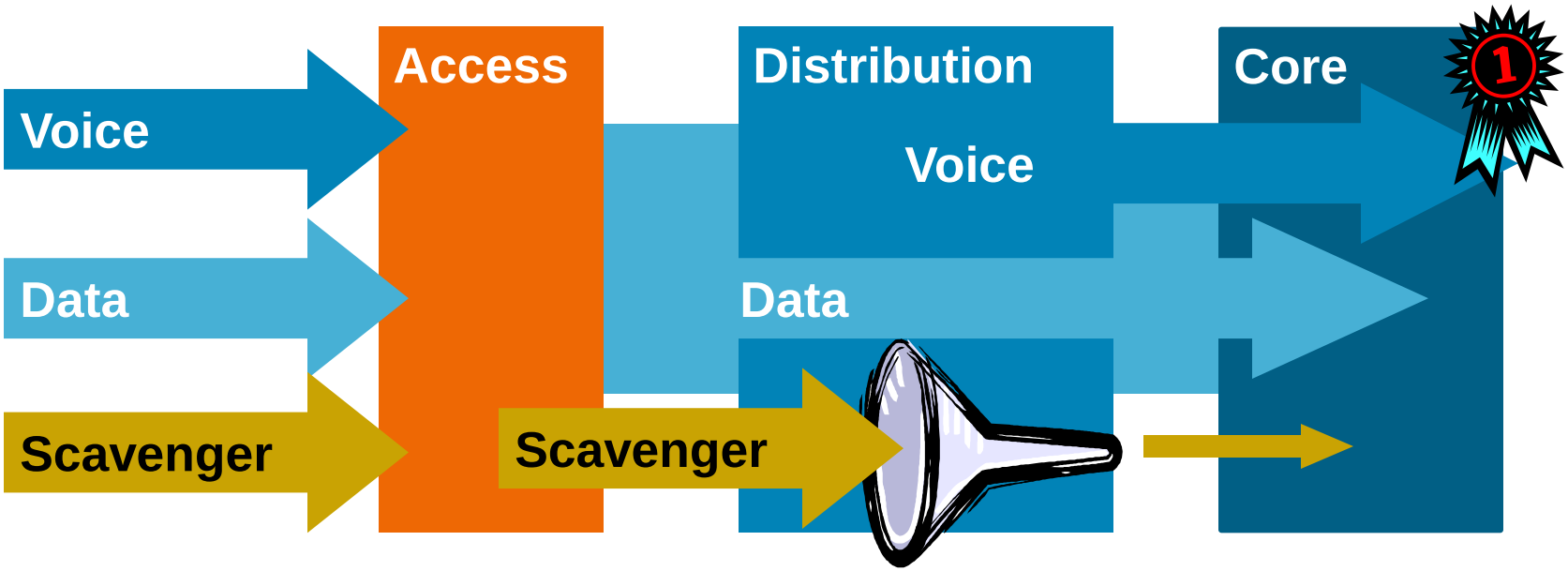


- Port security prevents CAM attacks and DHCP Starvation attacks
- DHCP Snooping prevents Rogue DHCP Server attacks
- Dynamic ARP Inspection prevents current ARP attacks
- IP Source Guard prevents IP/MAC Spoofing

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Use QoS Scavenger Class to Protect the Network

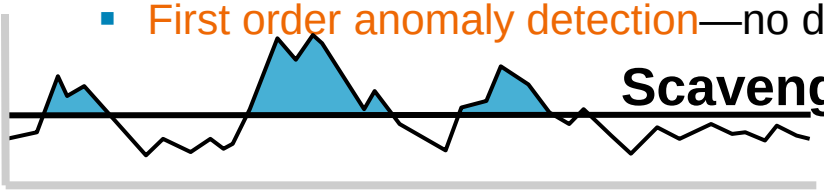
- Identify bad traffic and mark as scavenger class
- Scavenger class is an Internet-2 Draft Specification → CS1/CoS1



© 2010 Cisco and/or its affiliates. All rights reserved.

Identify Potential Worm Traffic and Drop It During Abnormal Conditions

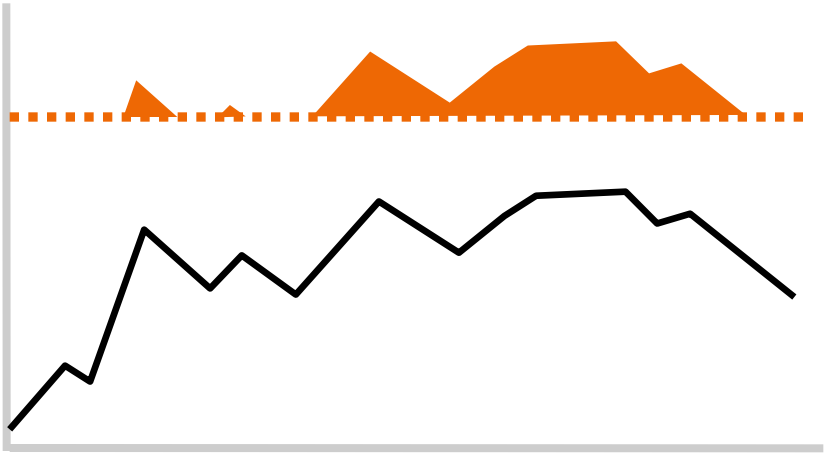
- All end systems generate traffic spikes
- Sustained traffic loads beyond ‘normal’ from each source device are considered suspect and marked as scavenger
- First order anomaly detection—no direct action taken



Network Entry Points

- During ‘abnormal’ worm traffic conditions traffic marked as Scavenger is aggressively dropped—second order detection
- Priority queuing ensuring low latency and minimum jitter for VoIP
- Stations not generating abnormal traffic volumes continue to receive good network service

- During ‘normal’ traffic conditions network is operating within designed capacity

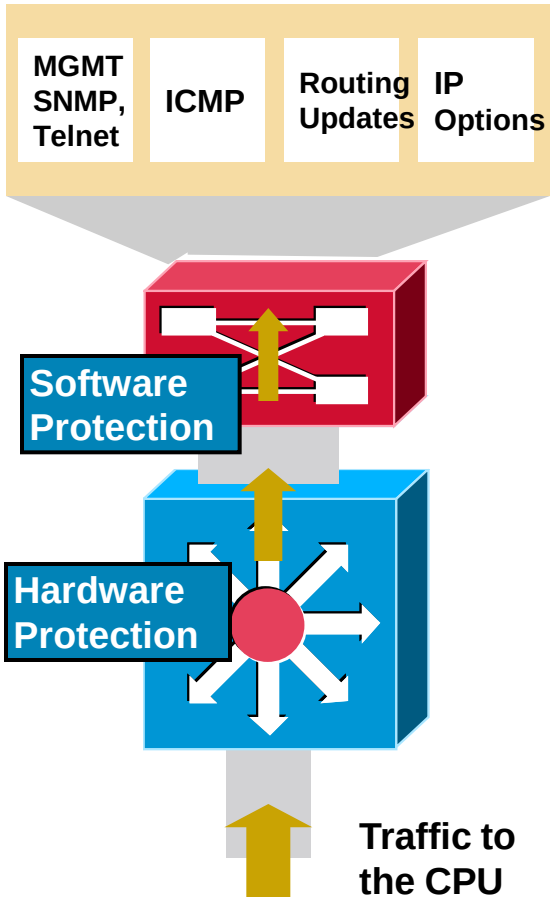


Aggregation Points

© 2010 Cisco and/or its affiliates. All rights reserved.

The Switches Also Need to Be Protected

- System CPU still has to be able to process certain traffic
  - BPDUs, CDP, EIGRP, OSPF
  - Telnet, SSH, SNMP
  - ARP, ICMP, IGMP
- Throttling on CPU-bound traffic helps protect the critical traffic
  - IOS Based SW Rate Limiters
  - Multiple CPU queues on 4500 & 3750
  - Hardware Rate Limiters on 6500 & Nexus 7000
  - Hardware Control Plane Policing (CoPP)



© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Summary

- **Optimize your network using a topology that give you the most options for network level redundancy**
- **Take a close look at Multi-Chassis Etherchannel solutions to really consolidate your network redundancy options into one**
- **Look at device level recovery options for speeding convergence in redundant topologies and absolutely leverage it at the access layer**
- **Don’t forget to protect your network from malicious activity**

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Q&A

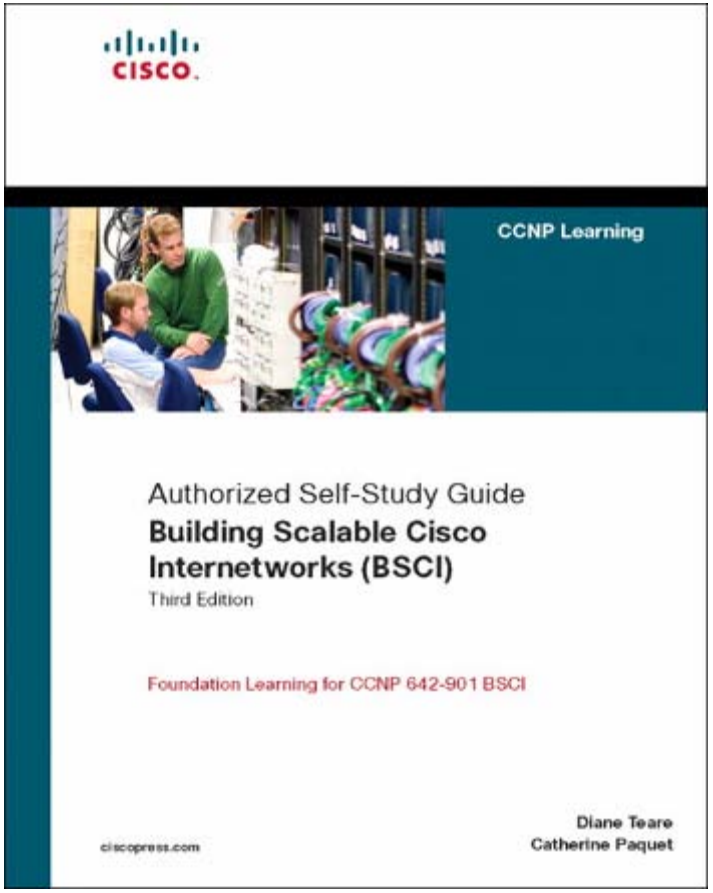
Recommended Reading

- Continue your Cisco Live learning experience with further reading from Cisco Press
- Check the Recommended Reading flyer for suggested books

  End-to-End QoS Network Design: Quality of Service in LANs, WANs, and VPNs
  ISBN: 1-58705-176-1

  Building Resilient IP Networks
  ISBN: 1-58705-215-6

  Top-Down Network Design, Second Ed.
  ISBN: 1-58705-152-4



Available Onsite at the Cisco Company Store

© 2010 Cisco and/or its affiliates. All rights reserved.

Complete Your Online Session Evaluation

- Give us your feedback and you could win fabulous prizes. Winners announced daily.
- Receive 20 Cisco Preferred Access points for each session evaluation you complete.
- Complete your session evaluation online now (open a browser through our wireless network to access our portal) or visit one of the Internet stations throughout the Convention Center.

Don’t forget to activate your Cisco Live and Networkers Virtual account for access to all session materials, communities, and on-demand and live activities throughout the year. Activate your account at any internet station or visit www.ciscolivevirtual.com.

© 2010 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

CISCO